

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:00:28 ; Search time 122.56 Seconds
(without alignments)
1.933 Million cell updates/sec

Title: US-08-653-294-1
Perfect score: 17
Sequence: 1 RXLXXXXXX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 9 | 52.9 | 6 | 1 W32886 | Peptide B2 major f |
| 2 | 9 | 52.9 | 6 | 1 W57229 | OA-519 peptide fra |
| 3 | 9 | 52.9 | 6 | 1 W84380 | Peptide B2 from a |
| 4 | 9 | 52.9 | 6 | 1 W96753 | Peptide fragment d |
| 5 | 9 | 52.9 | 7 | 1 R37426 | Promega peptide 1. |
| 6 | 9 | 52.9 | 7 | 1 R37432 | Promega peptide 7. |
| 7 | 9 | 52.9 | 7 | 1 R38351 | Phosphokinase subs |
| 8 | 9 | 52.9 | 7 | 1 R38352 | Phosphokinase subs |
| 9 | 9 | 52.9 | 7 | 1 R57736 | Protein-kinase inh |
| 10 | 9 | 52.9 | 7 | 1 R79663 | Protein kinase A p |
| 11 | 9 | 52.9 | 7 | 1 R83007 | Promega protein ki |
| 12 | 9 | 52.9 | 7 | 1 R94549 | Protein kinase sub |
| 13 | 9 | 52.9 | 7 | 1 R94550 | Protein kinase sub |
| 14 | 9 | 52.9 | 7 | 1 W09649 | Labelled peptide s |
| 15 | 9 | 52.9 | 7 | 1 W09655 | Labelled peptide s |
| 16 | 9 | 52.9 | 7 | 1 W53848 | Protein kinase sub |
| 17 | 9 | 52.9 | 7 | 1 W97426 | Shigella-like toxi |
| 18 | 9 | 52.9 | 8 | 1 P71338 | Sequence of novel |
| 19 | 9 | 52.9 | 8 | 1 R38355 | Phosphokinase subs |
| 20 | 9 | 52.9 | 8 | 1 R38356 | Phosphokinase subs |
| 21 | 9 | 52.9 | 8 | 1 R38357 | Phosphokinase subs |
| 22 | 9 | 52.9 | 8 | 1 R74539 | Protease Nexin-1 p |
| 23 | 9 | 52.9 | 9 | 1 R31108 | Synthetic sequence |
| 24 | 9 | 52.9 | 9 | 1 R79670 | Protein kinase A p |
| 25 | 9 | 52.9 | 9 | 1 W12322 | HINOS immunogenic |
| 26 | 9 | 52.9 | 9 | 1 W73276 | Influenza nucleopr |
| 27 | 9 | 52.9 | 10 | 1 R83095 | HLA-B2702 CTL modu |
| 28 | 9 | 52.9 | 10 | 1 R95426 | HLA-B2702.75-84(T) |
| 29 | 9 | 52.9 | 10 | 1 W21295 | Hydroxymethylgluta |
| 30 | 9 | 52.9 | 10 | 1 W07522 | T-cell modulating |
| 31 | 9 | 52.9 | 10 | 1 W07523 | T-cell modulating |
| 32 | 9 | 52.9 | 10 | 1 W11553 | Peptide NO54, Dete |
| 33 | 9 | 52.9 | 10 | 1 W12319 | Radiolabelled HINO |
| 34 | 9 | 52.9 | 10 | 1 W12320 | C-terminally subst |

35 9 52.9 10 1 W12321 C-terminally subst
36 9 52.9 10 1 W23651 Recombinant squirr
37 9 52.9 10 1 W23911 Human inducible ni
38 9 52.9 10 1 W33788 Peptide B2702.75-8
39 9 52.9 10 1 W33273 Influenza nucleopr
40 9 52.9 12 1 R86887 Human kidney thiope
41 9 52.9 12 1 W35792 Cucurbita maxima t
42 9 52.9 12 1 W35780 Cucurbita maxima t
43 9 52.9 12 1 W35787 Cucurbita maxima t
44 9 52.9 12 1 W35790 Cucurbita maxima t
45 9 52.9 12 1 W40716 Peptide which bind

ALIGNMENTS

RESULT 1

W32886
ID W32886 standard; Peptide; 6 AA.
AC W32886;
DT 16-JAN-1998 (first entry)
DE Peptide B2 major fragment of OA-519.
KW OA-519; cross-reaction; haptoglobin related; hpr; antibody;
KW epitope; haptoglobin 1; haptoglobin 2; cancer; breast cancer;
KW prognosis assay; peptide B2 major.
OS Homo sapiens.
PN US5665874-A.
PD 09-SEP-1997.
PF 17-JAN-1989; 297722.
PR 24-JAN-1984; US-188426.
PR 17-JAN-1989; US-297722.
PR 04-DEC-1990; US-622407.
PR 26-JUL-1991; US-735522.
PR 24-JUL-1992; US-917716.
PR 26-JUL-1993; US-096908.
PR 05-JUN-1995; US-469005.
PA (UIJO) UNIV JOHNS HOPKINS.
PI Kuhajda FP, Pasternack GR.
DR WPI; 97-469516/43.
PT DNA encoding protein cross-reactive with hpr gene product - useful
PT to raise antibodies reactive with epitope(s) found on hpr gene
PT product, useful in cancer, especially breast cancer, prognosis
PT assays
PS Example 12; Columns 41-42; 68pp; English.
CC The present sequence is a peptide fragment of a protein (OA-519)
CC cross-reactive with the haptoglobin related (hpr) gene product.
CC OA-519 can be used to raise antibodies reactive with epitopes.
CC found on the hpr gene product, but not on haptoglobin 1 or 2,
CC useful in cancer, especially breast cancer, prognosis assays.
SQ Sequence 6 AA;

Query Match 52.9%; Score 9; DB 1; Length 6;
Best Local Similarity 50.0%; Pred. No. 1.5e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
| |
Db 1 RAAL 4

RESULT 2

W57229
ID W57229 standard; peptide; 6 AA.
AC W57229;
DT 04-AUG-1998 (first entry)
DE OA-519 peptide fragment SEQ ID NO:5.
KW Human; hpr gene; haptoglobin; solid tumour; cancer; metastasis; breast;
KW screening; prognosis; haptoglobin; prostatic adenocarcinoma.
OS Synthetic.
OS Homo sapiens.
PN US579791-A.
PD 02-JUN-1998.

PF 24-JAN-1994; 188426.
 PR 24-JAN-1994; US-188426.
 PR 17-JAN-1989; US-297722.
 PR 04-DEC-1990; US-622407.
 PR 26-JUL-1991; US-735522.
 PR 21-JUL-1992; US-917716.
 PR 26-JUL-1993; US-096908.
 PA (UYJO) UNIV JOHNS HOPKINS.
 PI KuhaJda FP, Pasternack GR;
 DR WPI: 98-337128/29.
 PT Prognosis of solid tumours e.g. breast carcinoma - based on
 PT detection of OA-519 protein or mRNA in histological sections and
 PT biological fluids
 PS Example 12; Column 39; 50pp; English.
 CC The present sequence represents an OA-519 peptide fragment from an
 CC example of the present invention. The present invention describes a
 CC method and a kit for screening human samples to aid in determining
 CC the prognosis of breast carcinoma. The kit comprises in one or more
 CC containers: an antibody which specifically binds to one or more
 CC epitopes found on the sequence: Leu Tyr Ser Gly Asn Asp Val Thr Asp Ile
 CC Ser Asp Arg Phe Pro Lys Pro Glu Ile Ala Asn Gly Tyr Val Glu Lys
 CC Leu Phe Arg Tyr Glu Cys but not found on haptoglobin 1 or 2; and a
 CC reagent for detecting the antibody. The method comprises providing a
 CC histologic section from the tumour, and contacting the section with an
 CC antibody (as described above), and determining whether the antibody
 CC specifically binds to the section in cytoplasmic cellular regions, where
 CC the presence of the antibody binding correlates with a worsened
 CC prognosis of the solid tumour. The method is useful for the prognosis of
 CC solid tumours, especially breast carcinomas and prostatic
 CC adenocarcinomas.
 CC Sequence 6 AA;
 SQ

Query Match 52.9%; Score 9; DB 1; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 | |
 Db 1 RAAL 4

RESULT 3
 W84380
 ID W84380 standard; peptide; 6 AA.
 AC W84380;
 DT 26-MAR-1999 (first entry)
 DE Peptide B2 from a protein cross-reactive with the hpr gene product.
 KW Haptoglobin related gene; hpr gene; haptoglobin 1; haptoglobin 2;
 KW antibody; breast carcinoma; diagnostic marker; human solid tumour.
 OS Synthetic.
 OS Homo sapiens.
 PN US5864011-A.
 PD 26-JAN-1995.
 PF 05-JUN-1995; 469009.
 PR 24-JAN-1994; US-188426.
 PR 17-JAN-1989; US-297722.
 PR 04-DEC-1990; US-622407.
 PR 26-JUL-1991; US-735522.
 PR 24-JUL-1992; US-917716.
 PR 26-JUL-1993; US-096908.
 PR 05-JUN-1995; US-469009.
 PA (UYJO) UNIV JOHNS HOPKINS.
 PI KuhaJda FP, Pasternack GR;
 DR WPI: 99-131358/11.
 PT Polypeptide comprising epitope of hpr gene product - useful for
 PT producing antibodies for cancer prognosis
 PS Example 12; Column 25; 51pp; English.
 CC The present sequence represents a peptide derived from a protein
 CC that is immunologically cross-reactive with the haptoglobin related
 CC (hpr) gene product from breast cancer cells. The specification also
 CC describes a peptide that is immunologically cross-reactive with
 CC polyclonal antibodies that specifically bind the hpr gene product.

CC The peptide can be used to raise antibodies specific for a protein
 CC that is immunologically cross-reactive with the hpr gene product but
 CC not with haptoglobin 1 or haptoglobin 2. The peptide is found in the
 CC cytoplasm of breast carcinoma cells and is a useful diagnostic
 CC marker in human solid tumours for predicting the propensity for tumour
 CC invasion and early metastasis.
 SQ Sequence 6 AA;

Query Match 52.9%; Score 9; DB 1; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 | |
 Db 1 RAAL 4

RESULT 4
 W96753
 ID W96753 standard; peptide; 6 AA.
 AC W96753;
 DT 14-APR-1999 (first entry)
 DE Peptide fragment derived from protein OA-519.
 KW Monoclonal antibody; OA-519; haptoglobin 1; Hg1; Hg2;
 KW haptoglobin related protein; hpr; diagnostic marker; human solid tumour;
 KW tumour invasion; early metastasis; breast cancer.
 OS Homo sapiens.
 PN US5872217-A.
 PD 16-FEB-1999.
 PF 05-JUN-1995; 469007.
 PR 24-JAN-1994; US-188426.
 PR 17-JAN-1989; US-297722.
 PR 04-DEC-1990; US-622407.
 PR 26-JUL-1991; US-735522.
 PR 24-JUL-1992; US-917716.
 PR 26-JUL-1993; US-096908.
 PR 05-JUN-1995; US-469007.
 PA (UYJO) UNIV JOHNS HOPKINS.
 PI KuhaJda FP, Pasternack GR;
 DR WPI: 99-166717/14.
 PT New anti-OA-519 antibodies - useful as diagnostic and prognostic
 PT markers for human solid tumours, particularly breast cancers
 PS Example 12; Column 25; 51pp; English.
 CC Peptides W96750-53 are derived from the OA-519 protein. The
 CC specification describes monoclonal antibodies which specifically
 CC bind an epitope found on OA-519 but not on haptoglobin 1 (Hg1) or
 CC Hg2 or, optionally haptoglobin related protein (hpr) gene product.
 CC The antibodies can be used to detect OA-519 which is useful as a
 CC diagnostic marker in human solid tumours and for predicting the
 CC propensity for tumour invasion and early metastasis, particularly
 CC with breast cancers. They can detect aggressive tumour cells which
 CC result in decreased survival, increased metastasis, increased rates
 CC of clinical recurrence and overall worsened prognosis. The antibodies
 CC can also be used to purify haptoglobin related (hpr) proteins.
 SQ Sequence 6 AA;

Query Match 52.9%; Score 9; DB 1; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 | |
 Db 1 RAAL 4

RESULT 5
 R37426
 ID R37426 standard; peptide; 7 AA.
 AC R37426;
 DT 08-SEP-1993 (first entry)
 DE Promega peptide 1.

KW Modified peptide substrate; non-radioactive; detection: dansyl;
 KW sulphorhodamine 101; lissamine; rhodamine; enzymes; phosphatases;
 KW protein kinases; proteases.
 OS Synthetic.
 FH Key
 FT modified_site 1 /note= "detection tag= dansyl"
 FT Location/Qualifiers
 FN WO9310461-A.
 PD 27-MAY-1993.
 PF 12-NOV-1992; U09595.
 PR 12-NOV-1991; US-791928.
 PA (PROM-) PROMEGA CORP.
 PI Shultz JW, White DH;
 DR WPI; 93-182698/22.
 PT Quantitating presence or activity of enzyme - by incubating with
 PT modified peptide substrate and measuring the modified peptide
 PT prod.
 PS Claim 24; Page 27; 103pp; English.
 CC Promega peptide 1 is tagged with dansyl at residue 1 and may be used
 CC in a novel non-radioactive method of quantitating the presence or
 CC activity of an enzyme. The method can be used for rapid, specific
 CC and highly sensitive detection of enzymes such as protein kinases,
 CC phosphatases and proteases. They can be used to study enzyme
 CC function in metabolism and in diagnosis of disease. They also
 CC allow quantitative determ. of the enzyme's activity.
 CC See also R37427-36.
 SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RXXL 4
 | |
 Db 3 RASL 6

RESULT 6

ID R37432
 AC R37432;
 DT 08-SEP-1993 (first entry)
 DE Promega peptide 7.
 KW Modified peptide substrate; non-radioactive; detection: dansyl;
 KW sulphorhodamine 101; lissamine; rhodamine; enzymes; phosphatases;
 KW protein kinases; proteases.
 OS Synthetic.
 FH Key
 FT modified_site 1 /note= "detection tag= Lissamine. Rhodamine"
 FT Location/Qualifiers
 FN WO9310461-A.
 PD 27-MAY-1993.
 PF 12-NOV-1992; U09595.
 PR 12-NOV-1991; US-791928.
 PA (PROM-) PROMEGA CORP.
 PI Shultz JW, White DH;
 DR WPI; 93-182698/22.
 PT Quantitating presence or activity of enzyme - by incubating with
 PT modified peptide substrate and measuring the modified peptide
 PT prod.
 PS Claim 24; Page 27; 103pp; English.
 CC Promega peptide 7 is tagged with dansyl at residue 1 and may be used
 CC in a novel non-radioactive method of quantitating the presence or
 CC activity of an enzyme. The method can be used for rapid, specific
 CC and highly sensitive detection of enzymes such as protein kinases,
 CC phosphatases and proteases. They can be used to study enzyme
 CC function in metabolism and in diagnosis of disease. They also
 CC allow quantitative determ. of the enzyme's activity.
 CC See also R37426-36.
 SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 RXXL 4
 | |
 Db 3 RASL 6
 RESULT 7
 ID R38351
 AC R38351 standard; peptide; 7 AA.
 DT 15-OCT-1993 (first entry)
 DE Phosphokinase substrate peptide (1).
 KW Tumour-binding protein; target; substrate; phosphokinase; Kemptide.
 OS Synthetic.
 PN GB2262528-A.
 PD 23-JUN-1993.
 PF 15-DEC-1992; 026144.
 PR 16-DEC-1991; GB-026650.
 PA (BRTE-) BRITISH TECHNOLOGY GROUP LTD.
 PI Creighton AM;
 DR WPI; 93-199247/25.
 PT Structurally-modified tumour-binding proteins - used as
 PT targeting proteins, for treatment of tumours
 PS Claim 5; Page 13; 20pp; English.
 CC A protein that will bind to a tumour-associated structure has the
 CC amino gp. in at least one basic amino acid in the binding protein
 CC structurally modified to convert the amino gp., NH2, to the gp..
 CC -NH-CO-X-NHR
 CC R- H or an amino protecting gp.; and
 CC -CO-X-NHR- a residue of a peptide of formula NHR-X-COOH capable of
 CC acting as substrate for a phosphokinase.
 CC Examples of such substrate peptides include the heptapeptide that has
 CC become known as Kemptide having the sequence given in R38351 or related
 CC peptides given in R38352-63.
 SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RXXL 4
 | |
 Db 3 RASL 6

RESULT 8

ID R38352
 AC R38352 standard; peptide; 7 AA.
 DT 15-OCT-1993 (first entry)
 DE Phosphokinase substrate peptide (2).
 KW Tumour-binding protein; target; substrate; phosphokinase; Kemptide.
 OS Synthetic.
 PN GB2262528-A.
 PD 23-JUN-1993.
 PF 15-DEC-1992; 026144.
 PR 16-DEC-1991; GB-026650.
 PA (BRTE-) BRITISH TECHNOLOGY GROUP LTD.
 PI Creighton AM;
 DR WPI; 93-199247/25.
 PT Structurally-modified tumour-binding proteins - used as
 PT targeting proteins, for treatment of tumours
 PS Claim 5; Page 13; 20pp; English.
 CC A protein that will bind to a tumour-associated structure has the
 CC amino gp. in at least one basic amino acid in the binding protein
 CC structurally modified to convert the amino gp., NH2, to the gp..
 CC -NH-CO-X-NHR
 CC R- H or an amino protecting gp.; and
 CC -CO-X-NHR- a residue of a peptide of formula NHR-X-COOH capable of

CC acting as substrate for a phosphokinase.
 CC Examples of such substrate peptides include the heptapeptide that has
 CC become known as Kemptide having the sequence given in R38351 or related
 CC peptides given in R38352-63.
 SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 | |
 Db 3 RASL 6

RESULT 9

R55736
 ID R55736 standard; peptide; 7 AA.
 AC R55736;
 DT 16-NOV-1994 (first entry)
 DE Protein-kinase inhibitor.
 KW Protein-kinase inhibitor; fatty acyl-peptide; conjugate;
 KW antiproliferative; tumor; psoriasis; docosahexaenoic acid; DHA;
 KW eicosapentaenoic acid; EPA; antitumor.
 OS Synthetic.
 PN WO9412530-A.
 PD 09-JUN-1994.
 PF 29-NOV-1993; HU0065.
 PR 30-NOV-1992; US-984293.
 PA (BIOS-) BIOSIGNAL KUTATO FEJLESZTO KFT.
 PA (SYNT-) SYNTHETIC PEPTIDES INC.
 PI Balogh A, Cachia PJ, Hodges RS, Horvath A, Keri G;
 PI Szederkenyi F, Vadasz Z;
 WI: 94-200194/24.
 DR New fatty acyl-peptide conjugates for inhibiting cell
 PT proliferation - more active than free peptide, partic. for
 PT treating tumours, virus-infected cells, psoriasis, etc.
 PS Disclosure; Fig. 1; 45pp; English.
 CC The peptides given in R55718-48 can each be conjugated through an
 CC amide linkage with a polyunsaturated fatty acid moiety, such as
 CC docosahexaenoic acid or eicosapentaenoic acid, to improve
 CC antiproliferative activity. The cAMP-dependent
 CC protein-kinase inhibitor given in R55736 can be used to modulate
 CC native protein-kinases associated with cell proliferation.
 SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 | |
 Db 3 RASL 6

RESULT 10

R79663
 ID R79663 standard; peptide; 7 AA.
 AC R79663;
 DT 26-FEB-1996 (first entry)
 DE Protein kinase A phosphorylation site in Kemptide sequence.
 KW Peptide library; phosphorylation site; protein kinase; substrate;
 KW inhibitor; competitor; cellular response; cell cycle control;
 KW immune response; transcriptional activation; cell development.
 OS Synthetic.
 PN WO9518823-A2.
 PD 13-JUL-1995.
 PF 06-JAN-1995; U00147.
 PR 07-JAN-1994; US-178570.
 PA (BETH-) BETH ISRAEL HOSPITAL ASSOC.
 PI Cantley LC, Songyang Z;

DR WPI: 95-255036/33.
 PT Determ. of amino acid sequence of protein kinase phosphorylation
 PT site - by phosphorylation of peptide library and sequencing
 PT phosphopeptide(s) formed, also new substrates and their analogues
 PT for modulating or detecting protein kinase
 PS Example 5; Page 32; 131pp; English.
 CC An oriented degenerate peptide library of the amino acid formula
 CC R79661 was constructed to isolate the amino acid sequences at the
 CC phosphorylation sites of a protein kinase eg. protein kinase A,
 CC cyclin B/p33(cdc2), src family kinases, etc. Peptides which are
 CC phosphorylated are isolated and their amino acid sequences are compared
 CC to known substrate/inhibitor peptide sequences for that protein kinase.
 CC The peptides R79662-73 represent phosphorylation sites for protein
 CC kinase A. This peptide sequence is the phosphorylation site in a
 CC Kemptide sequence.
 CC The isolated peptides can be used to screen cpds. for effects on the
 CC protein kinase activity, generate antibodies to identify native kinase
 CC substrates, or modulate a variety of cellular responses in which protein
 CC kinases are involved eg. cell cycle control, immune response,
 CC transcriptional activation or cell development.
 SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 | |
 Db 3 RASL 6

RESULT 11

R83007
 ID R83007 standard; peptide; 7 AA.
 AC R83007;
 DT 11-APR-1996 (first entry)
 DE Promega protein kinase substrate peptide A.
 KW Promega; peptide; substrate; A; protein kinase; activity; assay.
 OS Synthetic.
 PN WO9523612-A1.
 PD 08-SEP-1995.
 PF 06-MAR-1995; U02856.
 PR 04-MAR-1994; US-208573.
 PA (PROM-) PROMEGA CORP.
 PI Goueli SA;
 WI: 95-320414/41.
 DR Quantitating activity of a selected protein kinase on a peptide
 PT substrate - by conjugating the substrate to a binding cpd.
 PS Claim 15; Page 33; 49pp; English.
 CC The activity of a selected protein kinase (PK), pref. Ser-Thr or
 CC Tyr PK, is quantified by conjugating a binding cpd. to a peptide
 CC substrate, pref. one of the Promega peptides A-H (R83007-14), adding
 CC this to a PK contg. soln., incubating the soln. to form a modified
 CC peptide prod. and then measuring the activity of the PK.
 SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 | |
 Db 3 RASL 6

RESULT 12

R94549
 ID R94549 standard; Peptide; 7 AA.
 AC R94549;
 DT 19-JUN-1996 (first entry)
 DE Protein kinase substrate peptide.

KW Cancer; therapy; radiotherapy; radiolabelling; phosphorus;
KW humanised antibody; complementarity determining region; CDR;
OS Synthesis;
PN WO9608506-A1.
PD 21-MAR-1996.
PF 18-SEP-1995; U11405.
PR 16-SEP-1994; US-308103.
PA (IMMU-) IMMUNOMEDICS INC.
PI Griffiths GL, Hansen HJ, Leung S;
DR WPI: 96-179897/18.
PT phosphorus labelled antibodies for use in cancer therapy - comprise
PT protein kinase peptide substrate sequence and specific targeting
PT CDR, for radio-labelling diseased cells with a phosphorus isotope
PS Claim 5; Page 43; 62pp; English.
CC Peptide substrates (R94549-56) for protein kinases are incorporated
CC into fusion proteins that also include an antibody complementarity
CC determining region. The fusion protein is produced by recombinant
CC DNA methods and can be expressed in prokaryotic or eukaryotic host
CC cells. The peptide substrate moiety of the fusion protein can be
CC radiolabelled by treatment with a protein kinase and a 32P- or
CC 33P-labelled phosphate donor. The radiolabelled fusion protein
CC is used for targeted radiotherapy of a patient suffering from a
CC tumour or an infectious lesion.
SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.5e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RXXL 4
| |
Db 3 RASL 6

RESULT 13

R94550
ID R94550 standard; Peptide; 7 AA.
AC R94550;
DE 19-JUN-1996 (first entry)
KW Protein kinase substrate peptide.
KW Cancer; therapy; radiotherapy; radiolabelling; phosphorus;
KW humanised antibody; complementarity determining region; CDR;
KW targeting; protein kinase.
OS Synthetic.
PN WO9608506-A1.
PD 21-MAR-1996.
PF 18-SEP-1995; U11405.
PR 16-SEP-1994; US-308103.
PA (IMMU-) IMMUNOMEDICS INC.
PI Griffiths GL, Hansen HJ, Leung S;
DR WPI: 96-179897/18.
PT phosphorus labelled antibodies for use in cancer therapy - comprise
PT protein kinase peptide substrate sequence and specific targeting
PT CDR, for radio-labelling diseased cells with a phosphorus isotope
PS Claim 5; Page 43; 62pp; English.
CC Peptide substrates (R94549-56) for protein kinases are incorporated
CC into fusion proteins that also include an antibody complementarity
CC determining region. The fusion protein is produced by recombinant
CC DNA methods and can be expressed in prokaryotic or eukaryotic host
CC cells. The peptide substrate moiety of the fusion protein can be
CC radiolabelled by treatment with a protein kinase and a 32P- or
CC 33P-labelled phosphate donor. The radiolabelled fusion protein
CC is used for targeted radiotherapy of a patient suffering from a
CC tumour or an infectious lesion.
SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.5e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RXXL 4
| |
Db 3 RASL 6

RESULT 14

W09649
ID W09649 standard; peptide; 7 AA.
AC W09649;
DE 20-MAY-1997 (first entry)
DE Labelled peptide substrate used in enzyme activity assay.
KW Enzyme activity; assay; measurement; label; rhodamine; dansyl;
KW non-radioactive; electrophoretic separation; protein kinase; protease;
KW phosphatase.
OS Synthetic.
FH Key Location/Qualifiers
FT modified_site 1 /note= "labelled with dansyl detection tag"
FT US5580747-A.
PN 03-DEC-1996.
PD 12-NOV-1991; 791928.
PF 12-NOV-1991; US-791928.
PR 21-JAN-1994; US-185448.
PA (PROM-) PROMEGA CORP.
PI Shultz JW, White DH;
DR WPI: 97-033588/03.
PT Non-radioactive assay for measuring enzyme activity - involving
PT electrophoretic sepn. of labelled cleavage prod. from labelled
PT peptide substrate
PS Claim 5; Column 37-38; 35pp; English.
CC W09649 is a peptide substrate used in a non-radioactive assay for
CC measuring enzyme activity. The assay comprises incubating the enzyme
CC with the labelled peptide substrate to form a labelled peptide
CC product; separating the product from the substrate by agarose gel
CC electrophoresis and measuring the amount of product by detecting the
CC label by fluorescence or chemiluminescence. The assay can be
CC performed rapidly and with great sensitivity. This peptide is
CC especially for determining cyclic AMP-dependent protein kinase
CC activity, e.g. to study its function in metabolism or to screen for
CC potential inhibitors.
SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.5e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RXXL 4
| |
Db 3 RASL 6

RESULT 15

W09655
ID W09655 standard; peptide; 7 AA.
AC W09655;
DE 20-MAY-1997 (first entry)
DE Labelled peptide substrate used in enzyme activity assay.
KW Enzyme activity; assay; measurement; label; rhodamine; dansyl;
KW non-radioactive; electrophoretic separation; protein kinase; protease;
KW phosphatase.
OS Synthetic.
FH Key Location/Qualifiers
FT modified_site 1 /note= "labelled with rhodamine B
FT detection tag"
FT US5580747-A.
PN 03-DEC-1996.
PD 12-NOV-1991; 791928.
PF 12-NOV-1991; US-791928.
PR 21-JAN-1994; US-185448.
PA (PROM-) PROMEGA CORP.
PI Shultz JW, White DH;

DR WPI: 97-033568/03.
PT Non-radioactive assay for measuring enzyme activity - involving
PT electrophoretic sepn. of labelled cleavage prod. from labelled
PT peptide substrate
PS Claim 5: Column 41-42: 35pp; English.
CC W09655 is a peptide substrate used in a non-radioactive assay for
CC measuring enzyme activity. The assay comprises incubating the enzyme
CC with the labelled peptide substrate to form a labelled peptide
CC product; separating the product from the substrate by agarose gel
CC electrophoresis and measuring the amount of product by detecting the
CC label by fluorescence or chemiluminescence. The assay can be
CC performed rapidly and with great sensitivity. This peptide is
CC especially for determining cyclic AMP-dependent protein kinase
CC activity, e.g. to study its function in metabolism or to screen for
CC potential inhibitors.
SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.5e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
| |
Db 3 RASL 6

Search completed: February 8, 2000, 01:29:33
Job time: 1745 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 05:08:50 ; Search time 117.7 Seconds
(without alignments)
4.008 Million cell updates/sec

Title: US-08-653-294-1
Perfect score: 17
Sequence: 1 RXXLXXXXX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : PIR_62:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----------|--------------------|
| 1 | 9 | 52.9 | 6 | 2 A11490 | Pyruvate kinase (E |
| 2 | 9 | 52.9 | 18 | 2 S54270 | GATA-2 protein - A |
| 3 | 9 | 52.9 | 18 | 2 S58277 | insulin-like growt |
| 4 | 9 | 52.9 | 20 | 2 T01691 | hypothetical prote |
| 5 | 9 | 52.9 | 20 | 2 A41717 | p100 protein - rat |
| 6 | 9 | 52.9 | 22 | 2 D30609 | Ig kappa chain V-I |
| 7 | 9 | 52.9 | 24 | 2 B30609 | Ig kappa chain V-I |
| 8 | 9 | 52.9 | 26 | 2 D30607 | Ig kappa chain V-I |
| 9 | 9 | 52.9 | 26 | 2 S12146 | hypothetical prote |
| 10 | 9 | 52.9 | 27 | 2 B4180 | phospholipase A2 (|
| 11 | 9 | 52.9 | 27 | 2 C30607 | Ig kappa chain V-I |
| 12 | 9 | 52.9 | 27 | 2 B61318 | momordin - balsam |
| 13 | 9 | 52.9 | 28 | 1 LFSEW | trip operon leader |
| 14 | 9 | 52.9 | 28 | 2 JQ0272 | hypothetical 3K pr |
| 15 | 9 | 52.9 | 30 | 2 S21816 | H+-transporting AT |
| 16 | 9 | 52.9 | 30 | 2 S21815 | H+-transporting AT |
| 17 | 9 | 52.9 | 30 | 2 I38914 | histidine--trNA li |
| 18 | 9 | 52.9 | 30 | 2 PH0882 | Ig kappa chain V r |
| 19 | 9 | 52.9 | 30 | 2 D70144 | hypothetical prote |
| 20 | 9 | 52.9 | 31 | 2 F30608 | Ig kappa chain V-I |
| 21 | 9 | 52.9 | 31 | 2 D30608 | Ig kappa chain V-I |
| 22 | 9 | 52.9 | 31 | 2 S22346 | transposase - Lact |
| 23 | 9 | 52.9 | 32 | 2 S22304 | hypothetical prote |
| 24 | 9 | 52.9 | 32 | 2 G42075 | finger protein (cl |
| 25 | 9 | 52.9 | 33 | 2 S22605 | ribosomal protein |
| 26 | 9 | 52.9 | 35 | 2 A61375 | basic fibroblast g |
| 27 | 9 | 52.9 | 37 | 2 A30607 | Ig kappa chain V-I |
| 28 | 9 | 52.9 | 38 | 1 F2R2L | photosystem II pro |
| 29 | 9 | 52.9 | 38 | 1 F2WTL | photosystem II pro |
| 30 | 9 | 52.9 | 38 | 1 F2SKL | photosystem II pro |

```

31 9 52.9 38 1 F2KTL photosystem II pro
32 9 52.9 38 2 T07254 photosystem II pro
33 9 52.9 38 2 S78339 photosystem II pro
34 9 52.9 38 2 S04064 photosystem II pro
35 9 52.9 38 2 S05685 photosystem II pro
36 9 52.9 38 2 S31821 photosystem II pro
37 9 52.9 38 2 A05048 photosystem II pro
38 9 52.9 38 2 S03193 photosystem II pro
39 9 52.9 38 2 S00691 photosystem II pro
40 9 52.9 38 2 S58566 photosystem II pro
41 9 52.9 38 2 C48310 photosystem II pro
42 9 52.9 38 2 S51366 photosystem II pro
43 9 52.9 38 2 S28056 photosystem II pro
44 9 52.9 38 2 S73310 photosystem II pro
45 9 52.9 38 2 T07477 photosystem II pro

```

ALIGNMENTS

RESULT 1

A11490
pyruvate kinase (EC 2.7.1.40) - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 03-Mar-1995
C:Accession: A11490

R:Hjeltnquist, G.; Andersson, J.; Edlund, B.; Engstrom, L.
Biochem. Biophys. Res. Commun. 61, 559-563, 1974
A:Title: Amino acid sequence of a (32-P)phosphopeptide from pig liver pyruvate kinase
A:Reference number: A11490; MUID:75127438

A:Accession: A11490

A:Molecule type: protein

A:Residues: 1-6 <HJE>

A:Experimental source: liver

C:Keywords: glycolysis; phosphotransferase

Query Match 52.9%; Score 9; DB 2; Length 6;

Best Local Similarity 50.0%; Pred. No. 1.4e+05;

Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4

Db 3 RASL 6

RESULT 2

S54270

GATA-2 protein - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 03-Nov-1995

C:Accession: S54270

R:Brewer, A.C.; Guille, M.J.; Fear, D.J.; Partington, G.A.; Patient, R.K.

EMBO J. 14, 757-766, 1995

A:Title: Nuclear translocation of a maternal CCAAT factor at the start of gastrulation

A:Reference number: S54270; MUID:95188880

A:Accession: S54270

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-18 <BRE>

QY 1 RXXL 4

Db 12 RAAL 15

Query Match 52.9%; Score 9; DB 2; Length 18;

Best Local Similarity 50.0%; Pred. No. 3.8e+02;

Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 3

S58277

insulin-like growth factor receptor type II - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 10-Sep-1997
C:Accession: S58277
R:Smrzka, O.W.; Stoger, R.; Kurzbauer, R.; Fae, I.; Fischer, G.F.; Barlow, D.P.
submitted to the EMBL Data Library, January 1995
A:Description: Conservation of a methylation imprint and a putative imprinting box at the
A:Reference number: S58277
A:Accession: S58277
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-18 <MR>
A:Cross-references: EMBL:X83702; NID:g929644; PID:g929645
C:Keywords: growth factor receptor

Query Match 52.9%; Score 9; DB 2; Length 18;
Best Local Similarity 50.0%; Pred. No. 3.8e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RXXL 4
| |
Db 8 RAAL 11

RESULT 4

T01691

hypotheical protein - maize chloroplast (fragment)

C:Species: Chloroplast Zea mays (maize)

C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 19-Feb-1999

C:Accession: T01691

R:Werr, W.

submitted to the EMBL Data Library, June 1982

A:Reference number: Z14397

A:Accession: T01691

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-20 <MR>

A:Cross-references: EMBL:V01469; NID:gl4307; PID:g388251

C:Genetics:

A:Genome: chloroplast

A:Note: Intron positions not resolved (incomplete sequence)

C:Keywords: chloroplast

Query Match 52.9%; Score 9; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 4.2e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RXXL 4
| |
Db 11 RSTL 14

RESULT 5

A41717

P100 protein - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 23-Mar-1993

C:Accession: A41717

R:Traub, L.M.; Sagi-Eisenberg, R.

J. Biol. Chem. 266, 24642-24649, 1991

A:Title: Purification of p100, a protein antigenically related to the signal transducing

A:Reference number: A41717; MUID:92105130

A:Accession: A41717

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-20 <RA>

Query Match 52.9%; Score 9; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 4.2e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RXXL 4
| |
Db 11 RSAL 14

RESULT 6

D30609

Ig kappa chain V-III regions (Jon and Mit) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 30-May-1997

C:Accession: D30609

R:Goni, F.R.; Chen, P.P.; McGinnis, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.; S

J. Immunol. 142, 3158-3163, 1989

A:Title: Structural and idiotypic characterization of the L chains of human IgM autoa

A:Reference number: A30601; MUID:89215279

A:Accession: D30609

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-22 <GON>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 52.9%; Score 9; DB 2; Length 22;
Best Local Similarity 50.0%; Pred. No. 4.6e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RXXL 4
| |
Db 18 RATL 21

RESULT 7

B30609

Ig kappa chain V-III region (She) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 30-May-1997

C:Accession: B30609

R:Goni, F.R.; Chen, P.P.; McGinnis, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.; S

J. Immunol. 142, 3158-3163, 1989

A:Title: Structural and idiotypic characterization of the L chains of human IgM autoa

A:Reference number: A30601; MUID:89215279

A:Accession: B30609

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-24 <GON>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 52.9%; Score 9; DB 2; Length 24;
Best Local Similarity 50.0%; Pred. No. 5e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RXXL 4
| |
Db 18 RATL 21

RESULT 8

D30607

Ig kappa chain V-III region (Sal) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 30-May-1997

C:Accession: D30607

R:Goni, F.R.; Chen, P.P.; McGinnis, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.; S

J. Immunol. 142, 3158-3163, 1989

A:Title: Structural and idiotypic characterization of the L chains of human IgM autoa

A:Reference number: A30601; MUID:89215279

A:Accession: D30607

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-26 <GON>
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotrimer; immunoglobulin

Query Match 52.9%; Score 9; DB 2; Length 26;
 Best Local Similarity 50.0%; Pred. No. 5.4e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 | |
 Db 18 RATL 21

RESULT 9
 S12146
 Hypothetical protein E16 - phage D108 (fragment)
 C:Species: phage D108
 C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Sep-1997
 C:Accession: S12146
 R:Pato, M.L.; Banerjee, M.; Wagonner, B.T.
 Nucleic Acids Res. 18, 6458, 1990
 A:Title: Sequence of gene E15 of bacteriophage D108 and comparison with phage Mu.
 A:Reference number: S12145; MUID:91057162
 A:Accession: S12146
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-26 <PAT>
 A:Cross-references: EMBL:X54298; NID:g14794; PID:g14796

Query Match 52.9%; Score 9; DB 2; Length 26;
 Best Local Similarity 50.0%; Pred. No. 5.4e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 | |
 Db 3 RYSL 6

RESULT 10
 B44180
 phospholipase A2 (EC 3.1.1.4) taicatoxin - Australian taipan (fragment)
 C:Species: Oxyuranus scutellatus scutellatus (Australian taipan)
 C:Date: 27-Apr-1993 #sequence_revision 07-Jun-1996 #text_change 07-Jun-1996
 C:Accession: S21101; B44180
 R:Possani, L.D.; Mochca-Morales, J.; Amezcua, J.; Martin, B.M.; Prestipino, G.; Nobile,
 Biochim. Biophys. Acta 1134, 210-216, 1992
 A:Title: Anionic currents of chick sensory neurons are affected by a phospholipase A(2)
 A:Reference number: S21101; MUID:92215788
 A:Accession: S21101
 A:Molecule type: protein
 A:Residues: 1-27 <POS>
 R:Possani, L.D.; Martin, B.M.; Yatani, A.; Mochca-Morales, J.; Zamudio, F.Z.; Gurrola, G.
 Toxicon 30, 1343-1364, 1992
 A:Title: Isolation and physiological characterization of taicatoxin, a complex toxin with
 A:Reference number: A44180; MUID:93134601
 A:Accession: B44180
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-27 <PO2>
 A:Experimental source: venom
 A:Note: sequence extracted from NCBI backbone (NCBIP:122481)
 C:Superfamily: phospholipase A2
 C:Keywords: carboxylic ester hydrolase; toxin

Query Match 52.9%; Score 9; DB 2; Length 27;
 Best Local Similarity 50.0%; Pred. No. 5.7e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 | |

Db 17 RSAL 20

RESULT 11

C30607
 Ig kappa chain V-III regions (Ang, Cin, and Sar) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 29-Jun-1989 #sequence_revision 01-Oct-1992 #text_change 30-May-1997
 C:Accession: C30607
 R:Goni, F.R.; Chen, P.P.; McGinnis, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.; S
 J. Immunol. 142, 3158-3163, 1989
 A:Title: Structural and idiotypic characterization of the L chains of human Igm autoa

A:Reference number: A30601; MUID:89215279
 A:Accession: C30607
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-27 <GON>
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotrimer; immunoglobulin

Query Match 52.9%; Score 9; DB 2; Length 27;
 Best Local Similarity 50.0%; Pred. No. 5.7e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 | |
 Db 18 RATL 21

RESULT 12

B61318
 momordin - balsam pear (fragment)
 C:Species: Momordica charantia (balsam pear, bitter gourd)
 C:Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 07-May-1999
 C:Accession: B61318
 R:Li, S.S.L.
 Experientia 36, 524-527, 1980
 A:Title: Purification and partial characterization of two lectins from Momordica char
 A:Reference number: A61318; MUID:80201763
 A:Accession: B61318
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-27 <LIA>
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
 C:Keywords: lectin; seed

Query Match 52.9%; Score 9; DB 2; Length 27;
 Best Local Similarity 50.0%; Pred. No. 5.7e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 | |
 Db 22 RATL 25

RESULT 13

LFESEW
 trp operon leader peptide - Serratia marcescens
 C:Species: Serratia marcescens
 C:Date: 18-Dec-1981 #sequence_revision 18-Dec-1981 #text_change 13-Nov-1998
 C:Accession: A03591
 R:Miozzari, G.F.; Yanofsky, C.
 Nature 276, 684-689, 1978
 A:Title: The regulatory region of the trp operon of Serratia marcescens.
 A:Reference number: A93202; MUID:79093989
 A:Accession: A03591
 A:Molecule type: DNA
 A:Residues: 1-28 <MIO>
 C:Genetics:
 A:Gene: trpL
 C:Function:

A:Description: involved in control of tryptophan operon transcription by attenuation
C:Superfamily: trp leader peptide

Query Match 52.9%; Score 9; DB 1; Length 28;
Best Local Similarity 50.0%; Pred. No. 5.9e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
| |
Db 21 RTSL 24

RESULT 14

JQ0272
hypothetical 3K protein (trnH-trnV intergenic region) - rice chloroplast
C:Species: chloroplast Oryza sativa (rice)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 20-Mar-1998
C:Accession: JQ0272; S05152
R:Shimada, H.; Whittier, R.F.; Hiratsuka, J.; Maeda, Y.; Hirai, A.; Sugiura, M.
submitted to JIPID, December 1989
A:Reference number: JQ0200
A:Accession: JQ0272
A:Molecule type: DNA
A:Residues: 1-28 <SHI>
A:Experimental source: cv. Nihonbare
R:Hiratsuka, J.; Shimada, H.; Whittier, R.; Ishibashi, T.; Sakamoto, M.; Mori, M.; Kondoh, M.
Mol. Gen. Genet. 217, 185-194, 1989
A:title: The complete sequence of the rice (Oryza sativa) chloroplast genome: intermediate of the cereals.
A:Reference number: S05080; MUID:89364698
A:Accession: S05152
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-28 <HIR>
A:Cross-references: GB:X15901; NID:g11957; PID:e286175; PID:g12032
A:Note: this sequence was submitted to EMBL, July 1989
C:Genetics:
A:Genome: chloroplast
C:keywords: chloroplast

Query Match 52.9%; Score 9; DB 2; Length 28;
Best Local Similarity 50.0%; Pred. No. 5.9e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
| |
Db 7 RSSL 10

RESULT 15

S21816
H+-transporting ATPase (EC 3.6.1.35) 70K chain - whisk fern (Psilotum nudum) (fragment)
C:Species: Psilotum nudum
C:Date: 25-Feb-1994 #sequence_revision 01-Mar-1996 #text_change 22-Jun-1999
C:Accession: S21816; S21817
R:Starke, T.; Linkkila, T.P.; Gogarten, J.P.
submitted to the EMBL Data Library, December 1990
A:Description: Two separate genes encode the catalytic 70kDa V-ATPase subunit in Psilotum
A:Reference number: S21814
A:Accession: S21816
A:Molecule type: DNA
A:Residues: 1-30 <STA>
A:Cross-references: EMBL:X56985; NID:g20604; PIDN:CAA40303.1; PID:e35787; PID:g1345563
A:Genetics: GEN1
A:Accession: S21817
A:Molecule type: DNA
A:Residues: 1-30 <STW>
A:Cross-references: EMBL:X56986; NID:g20606; PIDN:CAA40304.1; PID:e35789; PID:g1345564
A:Genetics: GEN2
C:Genetics: <GEN1>

A:Gene: P1
A:Introns: 2/2
C:Genetics: <GEN2>
A:Gene: P2

A:Introns: 2/2
C:Superfamily: vacuolar H+-transporting ATPase 69K chain; H+-transporting ATP synthase
C:Keywords: ATP; hydrolase

Query Match 52.9%; Score 9; DB 2; Length 30;
Best Local Similarity 50.0%; Pred. No. 6.3e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
| |
Db 25 RTTL 28

Search completed: February 7, 2000, 11:54:09
Job time: 24319 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 23:56:51 ; Search time 63.71 Seconds
(without alignments)
4.688 Million cell updates/sec

Title: US-08-653-294-1
Perfect score: 17
Sequence: 1 RXLXXXXXX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SwissProt_38.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|-------|--------|------------|--------|--------------|
| 1 | 9 | 52.9 | 24 | 1 | LPER_STRFR | P45440 | streptomyc |
| 2 | 9 | 52.9 | 26 | 1 | E16_BPD10 | P24796 | bacterioph |
| 3 | 9 | 52.9 | 28 | 1 | LPW_SERMA | P03055 | serratia ma |
| 4 | 9 | 52.9 | 30 | 1 | ENTF_SHIFL | P29698 | shigella fl |
| 5 | 9 | 52.9 | 30 | 1 | VAAI_PSIU | Q04237 | psilotum nu |
| 6 | 9 | 52.9 | 30 | 1 | VAA2_EQUAR | Q04238 | equisetum a |
| 7 | 9 | 52.9 | 30 | 1 | VAA2_PSIU | Q04239 | psilotum nu |
| 8 | 9 | 52.9 | 30 | 1 | Y357_BORBU | O51332 | borrelia bu |
| 9 | 9 | 52.9 | 30 | 1 | RL28_XENLA | P46780 | xenopus lae |
| 10 | 9 | 52.9 | 37 | 1 | PSBL_ANTMA | P29301 | antirrhinum |
| 11 | 9 | 52.9 | 37 | 1 | PSBL_ORISA | P12166 | oryza sativ |
| 12 | 9 | 52.9 | 38 | 1 | PSBL_CHLEU | P46306 | chlamydomon |
| 13 | 9 | 52.9 | 38 | 1 | PSBL_CHLRE | P32974 | chlamydomon |
| 14 | 9 | 52.9 | 38 | 1 | PSBL_CHLVU | P56339 | chlorella v |
| 15 | 9 | 52.9 | 38 | 1 | PSBL_CYPAP | P19154 | cyanophora |
| 16 | 9 | 52.9 | 38 | 1 | PSBL_EUGGR | P12228 | euglena gra |
| 17 | 9 | 52.9 | 38 | 1 | PSBL_CUIPH | O78464 | guillardia |
| 18 | 9 | 52.9 | 38 | 1 | PSBL_WARPO | P12165 | marichantia |
| 19 | 9 | 52.9 | 38 | 1 | PSBL_ODOSI | P49514 | odontella s |
| 20 | 9 | 52.9 | 38 | 1 | PSBL_FINTH | P41617 | pinus thunb |
| 21 | 9 | 52.9 | 38 | 1 | PSBL_PORPU | P51389 | porphyra pu |
| 22 | 9 | 52.9 | 39 | 1 | PSBL_SYNV3 | Q55354 | synecocyst |
| 23 | 9 | 52.9 | 44 | 1 | RIP3_MOMCH | P24817 | momordica c |
| 24 | 9 | 52.9 | 45 | 1 | RL34_STRBI | P25820 | streptomyc |
| 25 | 9 | 52.9 | 52 | 1 | FLMA_ECOLI | P16077 | escherichia |
| 26 | 9 | 52.9 | 52 | 1 | HOK_ECOLI | P11895 | escherichia |
| 27 | 9 | 52.9 | 52 | 1 | PLLA_CHICK | P26677 | gallus gall |
| 28 | 9 | 52.9 | 54 | 1 | LHB6_RHOAC | P35098 | rhodospseudo |
| 29 | 9 | 52.9 | 54 | 1 | LHB7_RHOAC | P35099 | rhodospseudo |
| 30 | 9 | 52.9 | 54 | 1 | RELX_BALED | P11185 | balaenopter |
| 31 | 9 | 52.9 | 55 | 1 | VNEM_PVXXC | P10466 | potato viru |
| 32 | 9 | 52.9 | 56 | 1 | ZN27_HUMAN | P17033 | homo sapien |
| 33 | 9 | 52.9 | 59 | 1 | COX7_YEAST | P10174 | saccharomyc |
| 34 | 9 | 52.9 | 59 | 1 | CYC_BACCL | P28332 | bacillus ca |

35 9 52.9 60 1 HMEN_LAMPL P31534 lampetra pl
36 9 52.9 60 1 MERC_PSEAE P04139 pseudomonas
37 9 52.9 60 1 MERC_SHIFL P04337 shigella fl
38 9 52.9 61 1 ATPE_YEAST P21306 saccharomyc
39 9 52.9 61 1 FLAW_KLEOX P56268 klebsiella
40 9 52.9 61 1 Y01H_BPT4 P39427 bacterioph
41 9 52.9 62 1 YM32_MARPO P38474 marichantia
42 9 52.9 65 1 LHB2_ECTHL P1696 ectothiorho
43 9 52.9 65 1 Y16G_BPT4 P39511 bacterioph
44 9 52.9 67 1 PSBH_CYPAP P48105 cyanophora
45 9 52.9 68 1 VGC_BPAL3 P31279 bacterioph

ALIGNMENTS

RESULT 1
LPER_STRFR STANDARD; PRT; 24 AA.
ID LP45440;
AC P45440;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE ERTHROMYCIN RESISTANCE LEADER PEPTIDE (23S RRNA METHYLASE LEADER PEPTIDE).
OS Streptomyces fradiae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetales; Streptomycetaceae; Streptomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88169508.
RA KAMIMIYA S., WEISBLUM B.;
RT "Translational attenuation control of ermSF, an inducible resistance determinant encoding rRNA N-methyltransferase from Streptomyces fradiae.";
RL J. Bacteriol. 170:1800-1811(1988).
CC -!- FUNCTION: THIS PEPTIDE IS INVOLVED IN THE CONTROL MECHANISM OF THE SYNTHESIS OF THE MACROLIDE-LINCOSAMIDE-STREPTOGRAMIN B RESISTANCE PROTEIN.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-3 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M19269; AAA26741.1; -;
DR Antibiotic resistance; Leader peptide.
KW SEQUENCE 24 AA; 2529 MW; 91CB30C9 CRC32;
SQ

Query Match 52.9%; Score 9; DB 1; Length 24;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
| |
Db 11 RAAL 14

RESULT 2
E16_BPD10 STANDARD; PRT; 26 AA.
ID E16_BPD10
AC P24796;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE PROTEIN E16 (FRAGMENT).
GN E16.
OS Bacteriophage D108.

OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91057162.
 RA PATO M.L., BANERJEE M., WAGONER B.T.;
 RT "Sequence of gene E15 of bacteriophage D108 and comparison with phage
 MU.";
 RL Nucleic Acids Res. 18:6458-6458(1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X54298; CAA38198.1; -;
 DR PIR; S12146; S12146. 26
 FT NON_TER 26
 SQ SEQUENCE 26 AA; 3102 MW; FCC4ED24 CRC32;

Query Match 52.9%; Score 9; DB 1; Length 26;
 Best Local Similarity 50.0%; Pred. No. 2.5e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 |
 Db 3 RTSL 6

RESULT 3
 LPW_SERMA
 ID LPW_SERMA STANDARD; PRT; 28 AA.
 AC P03055;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DE 01-MAY-1991 (Rel. 18, Last annotation update)
 DE TRP OPERON LEADER PEPTIDE.
 GN TRPL.
 OS Serratia marcescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Serratia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 79093989.
 RA MIOZZARI G.F., YANOFSKY C.;
 RT "The regulatory region of the trp operon of Serratia marcescens.";
 RL Nature 276:684-689(1978).
 CC -!- FUNCTION: THIS PROTEIN IS INVOLVED IN CONTROL OF THE BIOSYNTHESIS
 CC OF TRYPTOPHAN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J01791; -; NOT_ANNOTATED_CDS.
 DR PIR; A03591; LFSEW.
 KW Tryptophan biosynthesis; Leader peptide.
 SQ SEQUENCE 28 AA; 3391 MW; 3C9CA0F6 CRC32;

Query Match 52.9%; Score 9; DB 1; Length 28;
 Best Local Similarity 50.0%; Pred. No. 2.7e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 |
 Db 1

Db 21 RTSL 24
 RESULT 4
 ENTF_SHIFL
 ID ENTF_SHIFL STANDARD; PRT; 30 AA.
 AC P29698;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DE 15-DEC-1998 (Rel. 37, Last annotation update)
 DE ENTEROACTIN SYNTHETASE COMPONENT F (ENTEROCHELIN SYNTHASE F)
 DE (SERINE ACTIVATING ENZYME) (FRAGMENT).
 GN ENTF.
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Shigella.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91100373.
 RA SCHMITT M.P., PAYNE S.M.;
 RT "Genetic analysis of the enterobactin gene cluster in Shigella
 flexneri.";
 RL J. Bacteriol. 173:816-825(1991).
 CC -!- FUNCTION: ACTIVATION OF THE CARBOXYLATE GROUP OF L-SERINE VIA
 CC ATP-DEPENDENT PPI EXCHANGE REACTIONS TO THE AMINO ACYLADENYLATE,
 CC PREPARING THAT MOLECULE FOR THE FINAL STAGES OF ENTEROACTIN AND
 CC SYNTHESIS. ENTF MAY, IN THE PRESENCE OF 2,3-DIHYDROXYBENZOATE AND
 CC ENTE, GENERATE THE AMIDE BOND LINKAGE PRESENT 3 TIMES IN THE
 CC CYCLIC (2,3-DIHYDROXYBENZYL)SERINE TRIMER ENTEROACTIN.
 CC -!- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTHETHEINE.
 CC -!- PATHWAY: ENTEROACTIN BIOSYNTHESIS. ENTEROACTIN IS AN IRON-
 CC CHELATING COMPOUND INVOLVED IN TRANSPORTING IRON FROM THE
 CC BACTERIAL ENVIRONMENT INTO THE CELL CYTOPLASM.
 CC -!- SUBUNIT: PROTEINS ENTF, ENTF, ENTF, AND ENTF FORM A MULTIZYME
 CC COMPLEX CALLED ENTEROCHELIN SYNTHASE.
 CC -!- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M63304; -; NOT_ANNOTATED_CDS.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; PARTIAL.
 DR PROSITE; PS00455; AMP_BINDING; PARTIAL.
 KW Enterobactin biosynthesis; Iron transport; Phosphopantetheine;
 KW Ligase.
 FT NON_TER 1
 SQ SEQUENCE 30 AA; 3398 MW; 5FE5B79F CRC32;

Query Match 52.9%; Score 9; DB 1; Length 30;
 Best Local Similarity 50.0%; Pred. No. 2.9e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 |
 Db 25 RATL 28

RESULT 5
 VAAL_PSINU
 ID VAAL_PSINU STANDARD; PRT; 30 AA.
 AC Q04237;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE VACUOLAR ATP SYNTHASE CATALYTIC SUBUNIT A, ISOFORM 1 (EC 3.6.1.34)
 DE (FRAGMENT).
 DE

OS Psilotum nudum (Whisk fern).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Psilotophyta; Psilotales; Psilotaceae; Psilotum.
 RN [1]
 RX MEDLINE: 93138084.
 RA STARKE T., GOGARTEN J.P.;
 RT "A conserved intron in the V-ATPase A subunit genes of plants and
 RT algae.";
 RL FEBS Lett. 315:252-258(1993).
 CC -!- FUNCTION: CATALYTIC SUBUNIT OF THE PERIPHERAL V1 COMPLEX OF
 CC VACUOLAR ATPASE. V-ATPASE VACUOLAR ATPASE IS RESPONSIBLE FOR
 CC ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN EUKARYOTIC
 CC CELLS.
 CC -!- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A
 CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,
 CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE
 CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN).
 CC -!- MISCELLANEOUS: TWO SEPARATE GENES ENCODE THE CATALYTIC 70KDA
 CC V-ATPASE SUBUNIT IN PSILOSUM AND EQUISETUM.
 CC -!- SIMILARITY: TO THE BETA CHAINS OF ATP SYNTHASES AND TO A LESSER
 CC EXTENT TO THE ALPHA CHAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X56985; CAA40303.1; -;
 CC PROSITE: PS00152; ATPASE_ALPHA_BETA; PARTIAL.
 CC ATP synthesis; Hydrogen ion transport; Hydrolase; ATP-binding;
 CC Multigene family.
 CC FT NON_TER 1 1
 CC FT NON_TER 30 30
 CC SQ SEQUENCE 30 AA: 3380 MW: 55C24CFA CRC32;
 CC
 CC Query Match 52.9%; Score 9; DB 1; Length 30;
 CC Best Local Similarity 50.0%; Pred. No. 2.9e+02;
 CC Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC Qy 1 RXXL 4
 CC | |
 CC Db 25 RTTL 28
 CC
 CC RESULT 6
 CC VAA2_EQUAR VAA2_EQUAR STANDARD; PRT; 30 AA.
 CC AC Q04238;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 15-JUL-1998 (Rel. 36, Last annotation update)
 CC DE VACUOLAR ATP SYNTHASE CATALYTIC SUBUNIT A, ISOFORM 2 (EC 3.6.1.34)
 CC (FRAGMENT).
 CC OS Equisetum arvense (Field horsetail) (Common horsetail).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Equisetophyta; Sphenopsida; Equisetales; Equisetaceae;
 CC Equisetum.
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RX MEDLINE: 93138084.
 CC RA STARKE T., GOGARTEN J.P.;
 CC "A conserved intron in the V-ATPase A subunit genes of plants and
 CC algae.";
 CC RL FEBS Lett. 315:252-258(1993).
 CC -!- FUNCTION: CATALYTIC SUBUNIT OF THE PERIPHERAL V1 COMPLEX OF
 CC VACUOLAR ATPASE. V-ATPASE VACUOLAR ATPASE IS RESPONSIBLE FOR
 CC ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN EUKARYOTIC
 CC CELLS.

CC -!- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A
 CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,
 CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE
 CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN).
 CC -!- MISCELLANEOUS: TWO SEPARATE GENES ENCODE THE CATALYTIC 70KDA
 CC V-ATPASE SUBUNIT IN PSILOSUM AND EQUISETUM.
 CC -!- SIMILARITY: TO THE BETA CHAINS OF ATP SYNTHASES AND TO A LESSER
 CC EXTENT TO THE ALPHA CHAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X56984; CAA40302.1; -;
 CC PROSITE: PS00152; ATPASE_ALPHA_BETA; PARTIAL.
 CC ATP synthesis; Hydrogen ion transport; Hydrolase; ATP-binding;
 CC Multigene family.
 CC FT NON_TER 1 1
 CC FT NON_TER 30 30
 CC SQ SEQUENCE 30 AA: 3372 MW: BD406FF0 CRC32;
 CC
 CC Query Match 52.9%; Score 9; DB 1; Length 30;
 CC Best Local Similarity 50.0%; Pred. No. 2.9e+02;
 CC Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC Qy 1 RXXL 4
 CC | |
 CC Db 25 RTTL 28
 CC
 CC RESULT 7
 CC VAA2_PSINU VAA2_PSINU STANDARD; PRT; 30 AA.
 CC AC Q04239;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 15-JUL-1998 (Rel. 36, Last annotation update)
 CC DE VACUOLAR ATP SYNTHASE CATALYTIC SUBUNIT A, ISOFORM 2 (EC 3.6.1.34)
 CC (FRAGMENT).
 CC OS Psilotum nudum (Whisk fern).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Psilotophyta; Psilotales; Psilotaceae; Psilotum.
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RX MEDLINE: 93138084.
 CC RA STARKE T., GOGARTEN J.P.;
 CC "A conserved intron in the V-ATPase A subunit genes of plants and
 CC algae.";
 CC RL FEBS Lett. 315:252-258(1993).
 CC -!- FUNCTION: CATALYTIC SUBUNIT OF THE PERIPHERAL V1 COMPLEX OF
 CC VACUOLAR ATPASE. V-ATPASE VACUOLAR ATPASE IS RESPONSIBLE FOR
 CC ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN EUKARYOTIC
 CC CELLS.
 CC -!- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A
 CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,
 CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE
 CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN).
 CC -!- SIMILARITY: TO THE BETA CHAINS OF ATP SYNTHASES AND TO A LESSER
 CC EXTENT TO THE ALPHA CHAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

DR EMBL; X56986; CAA40304.1;
DR PROSITE; P500152; ATPASE_ALPHA_BETA; PARTIAL.
KW ATP synthesis; Hydrogen ion transport; Hydrolase; ATP-binding;
FT Multigene family.
FT NON_TER 1
FT NON_TER 30
FT NON_TER 30
SQ SEQUENCE 30 AA; 3380 MW; 55C24CFA CRC32;

Query Match 52.9%; Score 9; DB 1; Length 30;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
DB 25 RTTL 28

RESULT 8
Y357_BORBU
ID Y357_BORBU STANDARD; PRT; 30 AA.
AC 051332;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HYPOTHETICAL PROTEIN BB0357.
GN BB0357.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RX MEDLINE; 98065943.
RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,
RA VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
RA UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,
RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA SMITH H.O., VENTER J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X64209; CAB56812.1;
CC Ribosomal protein.
KW NON_TER 1
FT NON_TER 33
FT NON_TER 33
SQ SEQUENCE 33 AA; 3907 MW; 01329266 CRC32;

Query Match 52.9%; Score 9; DB 1; Length 33;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
DB 14 RSTL 17

RESULT 10
PSBL_ANTMA
ID PSBL_ANTMA STANDARD; PRT; 37 AA.
AC P29301;
DT 01-DEC-1992 (Rel. 24, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PHOTOSYSTEM II REACTION CENTER L PROTEIN (PSII 5 KD PROTEIN).
GN PSBL.
OS Antirrhinum majus (Garden snapdragon).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Asteridae; euasterids I; Lamiales; Scrophulariaceae;
OC Antirrhinum.
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-CV. SIPPE 50; TISSUE-LEAF;
RX MEDLINE; 92191997.
RA KUDLA J., IGOI G.L., METZLAFF M., HAGENMANN H., KOESSEL H.;
RT "RNA editing in tobacco chloroplasts leads to the formation of a
RT translatable psbL mRNA by a C to U substitution within the initiation
RT codon.";
RL EMBL J. 11:1099-1103(1992).
CC -!- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
CC B559.
CC -!- SIMILARITY: BELONGS TO THE PSBL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: X63206; CAA44889.1; -
 DR PIR: S19786; F2SKL; ANIMA:psbL;1.
 DR MENDEL: 2691; ANIMA:psbL;1.
 KW Photosynthesis; Photosystem II; Chloroplast.
 FT INIT_MET 0 BY SIMILARITY.
 SQ SEQUENCE 37 AA: 4339 MW: D9C3DE93 CRC32;

Query Match 52.9%; Score 9; DB 1; Length 37;
 Best Local Similarity 50.0%; Pred. No. 3.6e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 RXXL 4
 |
 Db 14 RTSL 17

RESULT 11
 PSBL_ORYSA STANDARD: PRT: 37 AA.
 AC P12166; P12167; Q34007;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PHOTOSYSTEM II REACTION CENTER L PROTEIN (PSII 5 KD PROTEIN).
 GN PSBL.
 OS Oryza sativa (Rice), Nicotiana tabacum (Common tobacco),
 OS Hordeum vulgare (Barley), Triticum aestivum (Wheat),
 OS Secale cereale (Rye), Zea mays (Maize), Pisum sativum (Garden pea),
 OS Spinacia oleracea (Spinach), Capsicum annuum (Bell pepper),
 OS Mesembryanthemum crystallinum (Common ice plant),
 OS Beta vulgaris (Sugar beet), and Populus deltoides (Poplar).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
 OC Poaceae; Oryza.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=O.SATIVA; STRAIN=CV. NIPPONBARE;
 RA SUGIURA M.;
 RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP COMPLETE GENOME.
 RC SPECIES=O.SATIVA;
 RX MEDLINE: 89364638.
 RA HIRATSUKA J., SHIMADA H., WHITTIER R., ISHIBASHI T., SAKAMOTO M.,
 RA MORI M., KONDO C., HONJI Y., SUN C.-R., MENG B.-Y., LI Y.-Q.,
 RA KANNO A., NISHIZAWA Y., HIRAI A., SHINOZAKI K., SUGIURA M.;
 RA "The complete sequence of the rice (Oryza sativa) chloroplast genome:
 RT intermolecular recombination between distinct trna genes accounts for
 RT a major plastid DNA inversion during the evolution of the cereals.";
 RL Mol. Gen. Genet. 217:185-194(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=N.TABACUM; STRAIN=CV. BRIGHT YELLOW 4;
 RA SUGIURA M.;
 RL Submitted (AUG-1986) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP COMPLETE GENOME.
 RC SPECIES=N.TABACUM;
 RA SHINOZAKI K., OHME M., TANAKA M., WAKASUGI T., HAYASHIDA N.,
 RA MATSUBAYASHI T., ZAITA N., CHUNWONGSE J., OBOKATA J.,
 RA YAMAGUCHI-SHINOZAKI K., OHTO C., TORAZAWA K., MENG B.-Y., SUGITA M.,
 RA DENO H., KAMOGASHIRA T., YAMADA K., KUSUDA J., TAKAIWA F., KATO A.,
 RA TOSHIOH N., SHIMADA H., SUGIURA M.;
 RA "The complete nucleotide sequence of the tobacco chloroplast genome:
 RT its gene organization and expression.";
 RL EMBO J. 5:2043-2049(1986).
 RN [5]

RNA EDITING OF INITIATOR CODON.
 RC SPECIES=N.TABACUM;
 RX CHAUDHURI S., MALIGA P.;
 RA "Sequences directing C to U editing of the plastid psbL mRNA are
 RT located within a 22 nucleotide segment spanning the editing site.";
 RL EMBO J. 15:5958-5964(1996).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC SPECIES=H.VULGARE; STRAIN=CV. SABARLIS;
 RX MEDLINE: 89240046.
 RA CHAKHMAKHICHEVA O.G., ANDREEVA A.V., BURYAKOVA A.A., REVERDATTO S.V.,
 RA EFIMOV V.A.;
 RA "Nucleotide sequence of the barley chloroplast psbE, psbF genes and
 RT flanking regions.";
 RL Nucleic Acids Res. 17:2858-2858(1989).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC SPECIES=H.VULGARE; STRAIN=CV. SABARLIS;
 RX MEDLINE: 92207253.
 RA EFIMOV V.A., ANDREEVA A.V., REVERDATTO S.V., CHAKHMAKHICHEVA O.G.;
 RA "Photosystem II of rye. Nucleotide sequence of the psbB, psbC, psbE,
 RT psbF, psbH genes of rye and chloroplast DNA regions adjacent to
 RT them.";
 RL Bioorg. Khim. 17:1369-1385(1991).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC SPECIES=T.AESTIVUM; STRAIN=CV. SENTRY; TISSUP-LEAF;
 RA WEBBER A.N., HIRD S.M., PACKMAN L.C., DYER T.A., GRAY J.C.;
 RA "A photosystem II polypeptide is encoded by an open reading frame
 RT co-transcribed with genes for cytochrome b-559 in wheat chloroplast
 RT DNA.";
 RL Plant Mol. Biol. 12:141-151(1989).
 RN [9]
 RP RNA EDITING OF INITIATOR CODON.
 RC SPECIES=T.AESTIVUM;
 RX MEDLINE: 92191997.
 RA KUDLA J., IGLOI G.L., METZLAFF M., HAGEMANN R., KOESSEL H.;
 RA "RNA editing in tobacco chloroplasts leads to the formation of a
 RT translatable psbL mRNA by a C to U substitution within the initiation
 RT codon.";
 RL EMBO J. 11:1099-1103(1992).
 RN [10]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.CEREALE;
 RX MEDLINE: 89160331.
 RA ZOLOTAREV A.S., KOLOSOV V.L.;
 RA "Nucleotide sequence of the rye chloroplast DNA fragment, comprising
 RT psbE and psbF genes.";
 RL Nucleic Acids Res. 17:1760-1760(1989).
 RN [11]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.CEREALE;
 RX MEDLINE: 90073796.
 RA KOLOSOV V.L., KLEZOVICH O.N., ABDULAEV N.G., ZOLOTAREV A.S.;
 RA "Photosystem II of rye. Nucleotide sequence of genes psbE, psbF, psbL
 RT and OPC40 of chloroplast DNA.";
 RL Bioorg. Khim. 15:1284-1286(1989).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC SPECIES=MAIZE;
 RA HALEY J., BOGORAD L.;
 RL Submitted (MAY-1989) to the EMBL/GenBank/DBJ databases.
 RN [13]
 RP SEQUENCE FROM N.A.
 RC SPECIES=MAIZE;
 RX MEDLINE: 95395841.
 RA MAIER R.M., NECKERMAN K., IGLOI G.L., KOESSEL H.;
 RA "Complete sequence of the maize chloroplast genome: gene content,
 RT hotspots of divergence and fine tuning of genetic information by
 RT transcript editing.";
 RL J. Mol. Biol. 251:614-628(1995).
 RN [14]

RP SEQUENCE FROM N.A.
 RC SPECIES-P. SATIVUM;
 RX MEDLINE: 89354671.
 RA WILLEY D.L., GRAY J.C.;
 RT "Two small open reading frames are co-transcribed with the pea
 RT chloroplast genes for the polypeptides of cytochrome b-559.";
 RL Curr. Genet. 15:213-220(1989).
 RN [15]
 RP SEQUENCE FROM N.A., AND RNA EDITING OF INITIATOR CODON.
 RC SPECIES-S. OLERACEA;
 RX MEDLINE: 93360903.
 RA BOCK R., HAGEMANN R., KOESSEL H., KUDLA J.;
 RT "Tissue- and stage-specific modulation of RNA editing of the psbF and
 RT psbL transcript from spinach plastids -- a new regulatory mechanism?";
 RL Mol. Gen. Genet. 240:238-244(1993).
 RN [16]
 RP SEQUENCE OF 1-12 FROM N.A.
 RC SPECIES-S. OLERACEA;
 RA HERMANN R.G., ALT J., SCHILLER B., WIDGER W.R., CRAWER W.A.;
 RT "Nucleotide sequence of the gene for apocytochrome b-559 on the
 RT spinach plastid chromosome: implications for the structure of the
 RT membrane protein.";
 RL FEBS Lett. 176:239-244(1984).
 RN [17]
 RP SEQUENCE FROM N.A., AND RNA EDITING OF INITIATOR CODON.
 RC SPECIES-C. ANNUM; STRAIN-CV. LAMUYO; TISSUE-LEAF, AND FRUIT;
 RX MEDLINE: 93099270.
 RA KUNTZ M., CAMARA B., WEIL J.-H., SCHANTZ R.;
 RT "The psbL gene from bell pepper (Capsicum annuum): plastid RNA
 RT editing also occurs in non-photosynthetic chromoplasts.";
 RL Plant Mol. Biol. 20:1185-1188(1992).
 RN [18]
 RP SEQUENCE FROM N.A.
 RC SPECIES-M. CRYSTALLINUM;
 RX MEDLINE: 94345017.
 RA FORTHOFEL N.R., CUSHMAN J.C.;
 RT "Characterization and expression of photosystem II genes (psbE, psbF,
 RT and psbL) from the facultative crassulacean acid metabolism plant
 RT Mesembryanthemum crystallinum.";
 RL Plant Physiol. 105:761-762(1994).
 RN [19]
 RP SEQUENCE FROM N.A.
 RC SPECIES-B. VULGARIS; STRAIN-CV. TK81-O; TISSUE-LEAF;
 RX MEDLINE: 95254673.
 RA KUBO T., YANAI Y., KINOSHITA T., MIKAMI T.;
 RT "The chloroplast trnT-trnW-petG gene cluster in the mitochondrial
 RT genomes of Beta vulgaris, B. trigyna and B. webbiana: evolutionary
 RT aspects.";
 RL Curr. Genet. 27:285-289(1995).
 RN [20]
 RP SEQUENCE FROM N.A.
 RC SPECIES-P. DELTOIDES; STRAIN-CV. STONEVILLE D121; TISSUE-LEAF;
 RA NAITHANI S.;
 RT Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [21]
 RP SEQUENCE OF 1-15 FROM N.A.
 RC SPECIES-T. AESTIVUM, AND S. OLERACEA;
 RX MEDLINE: 89121082.
 RA IKEUCHI M., TAKIO K., INOUE Y.;
 RT "N-terminal sequencing of photosystem II low-molecular-mass proteins.
 RT 5 and 4.1 kDa components of the O2-evolving core complex from higher
 RT plants.";
 RL FEBS Lett. 242:263-269(1989).
 RN [22]
 RP SEQUENCE OF 1-3, AND MASS-SPECTROMETRY.
 RC SPECIES-S. OLERACEA;
 RX MEDLINE: 98298118.
 RA ZHELEVA D., SHARMA J., PANICO M., MORRIS H.R., BARBER J.;
 RT "Isolation and characterization of monomeric and dimeric
 RT CP47-reaction center photosystem II complexes.";
 RL J. Biol. Chem. 273:16122-16127(1998).
 CC -!- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
 CC B559.

CC -!- MASS SPECTROMETRY: MW=4365.5; METHOD-MALDI.
 CC -!- SIMILARITY: BELONGS TO THE PSBL FAMILY.
 CC -!- CAUTION: THE INITIATOR METHIONINE IS CREATED BY RNA EDITING.
 CC -----
 Query Match 52.9%; Score 9; DB 1; Length 37;
 Best Local Similarity 50.0%; Pred. No. 3.6e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RXXL 4
 DB 14 RTSL 17
 RESULT 12
 PSBL_CHLEU STANDARD; PRT; 38 AA.
 ID PSBL_CHLEU
 AC P46306;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE PHOTOSYSTEM II REACTION CENTER L PROTEIN (PSII 5 KD PROTEIN).
 GN PSBL.
 OS Chlamydomonas eugametos.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadaceae; Chlamydomonas.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TURMEL M., OTIS C.;
 RL Submitted (XXX-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
 CC B559.
 CC -----
 -!- SIMILARITY: BELONGS TO THE PSBL FAMILY.
 CC -----
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L29282; AAA84157.1; -.
 DR MENDEL: 2694; CHLEU:psbl;1.
 KW PHOTOSYNTHESIS; PHOTOSYSTEM II; CHLOROPLAST.
 SQ SEQUENCE 38 AA; 4389 MW; EEC7F31 CRC32;
 Query Match 52.9%; Score 9; DB 1; Length 38;
 Best Local Similarity 50.0%; Pred. No. 3.7e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RXXL 4
 DB 15 RTSL 18
 RESULT 13
 PSBL_CHLEU STANDARD; PRT; 38 AA.
 ID PSBL_CHLEU
 AC P32974;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE PHOTOSYSTEM II REACTION CENTER L PROTEIN (PSII 5 KD PROTEIN).
 GN PSBL.
 OS Chlamydomonas reinhardtii.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadaceae; Chlamydomonas.
 RN [1]
 RP SEQUENCE FROM N.A.

```
RX MEDLINE: 92315354.
RA FONG S.E., SURZYCKI S.J.:
RT "Organization and structure of plastome psbF, psbL, petG and ORF712
RL genes in Chlamydomonas reinhardtii."
RL Curr. Genet. 21:527-530(1992).
CC -!- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
CC B559.
CC
CC -!- SIMILARITY: BELONGS TO THE PSBL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X66250; CAA46978.1;
DR PIR: S26879; S26879.
DR MENDEL: 2695; Chlure:psbl.1.
KW Photosynthesis; Photosystem II; Chloroplast.
SQ SEQUENCE 38 AA; 4428 MW; 2A6CC463 CRC32;

Query Match 52.9%; Score 9; DB 1; Length 38;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
DB 15 RTSL 18

RESULT 14
PSBL_CHLVU STANDARD; PRT; 38 AA.
ID PSBL_CHLVU
AC P56339;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PHOTOSYSTEM II REACTION CENTER L PROTEIN (PSII 5 KD PROTEIN).
GN PSBL.
OS Chlorella vulgaris.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Chlorella.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IAM C-27 / TAMIYA;
RX MEDLINE: 97303241.
RA WAKASUGI T., NAGAI T., KAPOOR M., SUGITA M., ITO M., ITO S.,
RA TSUDZUKI J., NAKASHIMA K., TSUDZUKI T., SUZUKI Y., HAMADA A., OHTA T.,
RA INAMURA A., YOSHINAGA K., SUGIURA M.;
RT "Complete nucleotide sequence of the chloroplast genome from the
RT green alga Chlorella vulgaris: the existence of genes possibly
RT involved in chloroplast division."
RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
CC -!- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
CC B559.
CC -!- SIMILARITY: BELONGS TO THE PSBL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB001684; BAA20643.1;
KW Photosynthesis; Photosystem II; Chloroplast.
SQ SEQUENCE 38 AA; 4388 MW; 1A947871 CRC32;
```

```
Query Match 52.9%; Score 9; DB 1; Length 38;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
DB 15 RTSL 18

RESULT 15
PSBL_CYPAA STANDARD; PRT; 38 AA.
ID PSBL_CYPAA
AC P19154;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PHOTOSYSTEM II REACTION CENTER L PROTEIN (PSII 5 KD PROTEIN).
GN PSBL.
OS Cyanophora paradoxa.
OC Cyanelle.
OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LB555 / PRINGSHEIM;
RA CANTRELL A., BRYANT D.A.;
RT "Nucleotide sequence of the genes encoding cytochrome b-559 from the
RT cyanelle genome of Cyanophora paradoxa."
RL Photosyn. Res. 16:65-81(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-LB555 / PRINGSHEIM;
RA CANTRELL A., BRYANT D.A.;
RT "Molecular cloning and nucleotide sequences of the genes encoding
RT cytochrome b-559 from the cyanelle genome of Cyanophora paradoxa."
RL Prog. Photosyn. Res. 4:659-662(1987).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-LB555 / PRINGSHEIM;
RA STREWALT V.L., MICHALOWSKI C.B., LUFFELHARDT W., BOHNET H.J.,
RA BRYANT D.A.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
CC B559.
CC -!- SIMILARITY: BELONGS TO THE PSBL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M35129; AAA31697.1;
DR EMBL: U30821; AAA81211.1;
DR PIR: S09483; F2KTL.
DR MENDEL: 7894; CYAPA:psbl.1.
KW Photosynthesis; Photosystem II; Cyanelle.
SQ SEQUENCE 38 AA; 4473 MW; 4277B4E3 CRC32;

Query Match 52.9%; Score 9; DB 1; Length 38;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
DB 15 RTSL 18
```

Search completed: February 8, 2000, 00:59:40

Job time: 3769 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 04:16:11 ; Search time 209.03 Seconds
(without alignments)
3.317 Million cell updates/sec

Title: US-08-653-294-1

Perfect score: 17

Sequence: 1 RXXLXXXXXX 10

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL.12.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------------|
| 1 | 9 | 52.9 | 9 | 11 | Q61723 mus musculus |
| 2 | 9 | 52.9 | 10 | 12 | Q69347 human herpe |
| 3 | 9 | 52.9 | 12 | 2 | Q47251 Q47251 escherichia |
| 4 | 9 | 52.9 | 19 | 10 | Q40183 lemna gibba |
| 5 | 9 | 52.9 | 20 | 2 | P72423 saccharopol |
| 6 | 9 | 52.9 | 20 | 8 | Q33294 sea mayas (m |
| 7 | 9 | 52.9 | 21 | 10 | Q40181 lemna gibba |
| 8 | 9 | 52.9 | 22 | 2 | Q46322 escherichia |
| 9 | 9 | 52.9 | 22 | 10 | Q9XG00 oryza sativ |
| 10 | 9 | 52.9 | 27 | 5 | Q61678 drosophila |
| 11 | 9 | 52.9 | 27 | 12 | Q92079 hepatitis c |
| 12 | 9 | 52.9 | 27 | 12 | Q92063 hepatitis c |
| 13 | 9 | 52.9 | 27 | 12 | Q92058 hepatitis c |
| 14 | 9 | 52.9 | 28 | 8 | Q37005 oryza sativ |
| 15 | 9 | 52.9 | 32 | 2 | Q00491 streptomyce |
| 16 | 9 | 52.9 | 33 | 11 | Q9WUX2 rattus norv |
| 17 | 9 | 52.9 | 34 | 5 | Q17148 echinococcu |
| 18 | 9 | 52.9 | 37 | 4 | Q9Y6H2 homo sapien |
| 19 | 9 | 52.9 | 37 | 8 | Q9XQ05 toxoplasma |
| 20 | 9 | 52.9 | 38 | 2 | Q47883 frankia aln |

| | | | | | |
|----|---|------|----|----|--------|
| 21 | 9 | 52.9 | 38 | 8 | Q47030 |
| 22 | 9 | 52.9 | 39 | 2 | Q86199 |
| 23 | 9 | 52.9 | 40 | 8 | Q32946 |
| 24 | 9 | 52.9 | 40 | 8 | Q32971 |
| 25 | 9 | 52.9 | 40 | 11 | Q922L3 |
| 26 | 9 | 52.9 | 42 | 2 | Q88038 |
| 27 | 9 | 52.9 | 42 | 5 | Q28253 |
| 28 | 9 | 52.9 | 42 | 8 | Q95868 |
| 29 | 9 | 52.9 | 42 | 12 | Q98643 |
| 30 | 9 | 52.9 | 42 | 12 | Q98648 |
| 31 | 9 | 52.9 | 42 | 12 | Q56133 |
| 32 | 9 | 52.9 | 42 | 12 | Q91867 |
| 33 | 9 | 52.9 | 42 | 12 | Q91870 |
| 34 | 9 | 52.9 | 43 | 5 | Q9XY96 |
| 35 | 9 | 52.9 | 43 | 8 | Q20131 |
| 36 | 9 | 52.9 | 43 | 12 | Q98646 |
| 37 | 9 | 52.9 | 44 | 5 | Q17881 |
| 38 | 9 | 52.9 | 44 | 8 | Q92073 |
| 39 | 9 | 52.9 | 44 | 12 | Q68313 |
| 40 | 9 | 52.9 | 44 | 12 | Q68326 |
| 41 | 9 | 52.9 | 44 | 12 | Q87847 |
| 42 | 9 | 52.9 | 45 | 12 | Q64828 |
| 43 | 9 | 52.9 | 45 | 12 | Q92289 |
| 44 | 9 | 52.9 | 45 | 12 | Q64838 |
| 45 | 9 | 52.9 | 46 | 2 | Q50275 |

ALIGNMENTS

RESULT 1

Q61723 ID Q61723 PRELIMINARY; PRT; 9 AA.
AC Q61723; 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-NOV-1996 (TREMREL. 12, Last annotation update)
DE NUCLEAR FACTOR OF KAPPA LIGHT CHAIN PROTEIN ENHANCER IN B-CELLS 1,
GN P105 (NF-KAPPA-B DNA-BINDING SUBUNIT) (FRAGMENT).
DE NFKB1 OR NF-KAPPA-B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/CBYJ; TISSUE=SPLEEN;
RX MEDLINE; 90367113.
RA GHOSH S., GIFFORD A.M., RIVIERE L.R., TEMPEST P., NOLAN G.P.,
RA BALTIMORE D.;
RT "Cloning of the p50 DNA binding subunit of NF-kappa B: homology to rel
and dorsal".
RL Cell 62:1019-1029(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/CBYJ; TISSUE=SPLEEN;
RX MEDLINE; 94156215.
RA HUO L., CHUNG W.H., ROTHSTEIN T.L.;
RT "C-terminal sequence of the NF-kappa B p50 precursor from primary
murine B-lymphocytes".
RL Gene 139:287-288(1994).
DR EMBL; L13466; AAC37644.1; .
DR MGD; MGI:97312; Nfkb1.
KW DNA-binding.
FT NON_TER 1 1
FT CONFLICT 5 5 A -> P (IN REF. 1).
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 925 MW; C5EBD462 CRC32;

Query Match 52.9%; Score 9; DB 11; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.3e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
|
DB 3 RTAL 6

RESULT 2

Q69347 PRELIMINARY; PRT; 10 AA.
ID Q69347;
AC Q69347;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE HYPOTHETICAL I.I KD PROTEIN.
OS human herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KOS;
RX MEDLINE; 86068025.
RA GIBBS J.S., CHIOU H.C., HALL J.D., MOUNT D.W., RETONDO M.J.,
RA WELER S.K., COEN D.M.;
RT "Sequence and mapping analyses of the herpes simplex virus DNA
RT polymerase gene predict a C-terminal substrate binding domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:7969-7973(1985).
DR EMBL; M10792; AAA66437.1; -.
KW Hypothetical protein.
SQ SEQUENCE 10 AA; 1057 MW; 223ABA1A CRC32;

Query Match

Best Local Similarity 52.9%; Score 9; DB 12; Length 10;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
|
DB 4 RASL 7

RESULT 3

Q47251 PRELIMINARY; PRT; 12 AA.
ID Q47251;
AC Q47251;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE HEMC GENE PRODUCT.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE; 88096587.
RA JORDAN P.M., MCBEE B.I.A., ALWAN A.F., THOMAS S.D.;
RT "Nucleotide sequence of hemd, the second gene in the hem operon of
RT Escherichia coli K-12.";
RL Nucleic Acids Res. 15:10583-10583(1987).
DR EMBL; Y00883; CAA68775.1; -.
SQ SEQUENCE 12 AA; 1375 MW; A41CD7AF CRC32;

Query Match

Best Local Similarity 52.9%; Score 9; DB 2; Length 12;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
|
DB 1 RSSL 4

RESULT 4

Q40183

ID Q40183 PRELIMINARY; PRT; 19 AA.
AC Q40183;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TRENBLrel. 01, Last annotation update)
DE NEGATIVELY LIGHT-REGULATED PROTEIN (LG106) (FRAGMENT).
OS Lemna gibba (Swollen duckweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Araceae;
OC Lemna.
RN [1]
RP SEQUENCE FROM N.A.
RA OKUBARA P.A., FLORES S., TOBIN E.M.;
RL Plant Mol. Biol. 11:673-681(1988).
DR EMBL; M35866; AAA3395.1; -.
FT NON_TER 1
SQ SEQUENCE 19 AA; 2153 MW; 76EELF6D CRC32;

Query Match

Best Local Similarity 52.9%; Score 9; DB 10; Length 19;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
|
DB 16 RASL 19

RESULT 5

P72423 PRELIMINARY; PRT; 20 AA.
ID P72423;
AC P72423;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE 6-DEOXYERYTHRONOLIDE B SYNTHASE (FRAGMENT).
GN ERYA.
OS Saccharopolyspora erythraea (Streptomyces erythraeus).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Pseudonocardineae; Pseudonocardaceae;
OC Saccharopolyspora.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93231529.
RA DONADIO S., STAVER M.J.;
RT "IS1136, an insertion element in the erythromycin gene cluster of
RT Saccharopolyspora erythraea.";
RL Gene 126:147-151(1993).
DR EMBL; L07626; AAA26506.1; -.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2339 MW; FB210B56 CRC32;

Query Match

Best Local Similarity 52.9%; Score 9; DB 2; Length 20;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
|
DB 13 RATL 16

RESULT 6

Q33294 PRELIMINARY; PRT; 20 AA.
ID Q33294;
AC Q33294;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TRENBLrel. 01, Last annotation update)
DE EITHER FRAME IS POSSIBLE (FRAGMENT).
OS Zea mays (Maize).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
OC Poaceae; Zea.

RN [1]

RP SEQUENCE FROM N.A.

RA WERR W.;

RL Submitted (JUN-1982) to the EMBL/GenBank/DBJ databases.

DR EMBL; V01469; CRA24716.1; -

KW Chloroplast.

FT NON_TER 1 1

FT NON_TER 20 20

SQ SEQUENCE 20 AA; 2377 MW; C25D1D4C CRC32;

Query Match 52.9%; Score 9; DB 8; Length 20;
Best Local Similarity 50.0%; Pred. No. 9.6e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RXXL 4

Db 11 RSTL 14

RESULT 7

Q40181

ID Q40181 PRELIMINARY; PRT; 21 AA.

AC Q40181;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-NOV-1996 (TREMBlrel. 01, Last annotation update)

DE SHORTEST ORF (FRAGMENT).

OS Lenna gibba (Swollen duckweed).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Araceae;

OC Lenna.

OC [1]

RP SEQUENCE FROM N.A.

RC STRAIN=L;

RA OKUBARA P.A.; FLORES S.; TOBIN E.M.;

RL Plant Mol. Biol. 11:673-681(1988).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=L;

RA OKUBARA P.A.; FLORES S.; TOBIN E.M.;

RL Plant Mol. Biol. 15:955-956(1990).

DR EMBL; X14075; CAA32238.1; -

FT NON_TER 1 1

SQ SEQUENCE 21 AA; 2324 MW; 9FC86BEE CRC32;

Query Match 52.9%; Score 9; DB 10; Length 21;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RXXL 4

Db 17 RASL 20

RESULT 8

ID Q54632

AC Q54632 PRELIMINARY; PRT; 22 AA.

DT 01-JUN-1998 (TREMBlrel. 06, Created)

DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)

DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE CLONING VECTOR PK1194T (FRAGMENT).

GN FUSED-CCDB.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 98172738.

RA GABANT P., SPIRER C.Y., COUTURIER M., FAELAN M.;

RT "Direct selection cloning vectors adapted to the genetic analysis of

RL gram-negative bacteria and their plasmids.";

RL Gene 207:87-92(1998).

DR EMBL; Y10544; CAA71573.1; -

DR EMBL; Y10543; CAA71571.1; -

FT NON_TER 22 22

SQ SEQUENCE 22 AA; 2433 MW; C397C596 CRC32;

Query Match 52.9%; Score 9; DB 2; Length 22;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RXXL 4

Db 11 RSTL 14

RESULT 9

Q9XGQ0

ID Q9XGQ0 PRELIMINARY; PRT; 22 AA.

AC Q9XGQ0;

DT 01-NOV-1999 (TREMBlrel. 12, Created)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)

DE HYPOTHETICAL PROTEIN.

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;

OC Poaceae; Oryza.

OC [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. NIPPONBARE;

RA SASAKI T., MATSUMOTO T., YAMAMOTO K.;

RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, PAC

RT clone:PO026F07.";

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP000364; BAA81770.1; -

SQ SEQUENCE 22 AA; 2500 MW; 6EE798D7 CRC32;

Query Match 52.9%; Score 9; DB 10; Length 22;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RXXL 4

Db 14 RASL 17

RESULT 10

Q61678

ID Q61678 PRELIMINARY; PRT; 27 AA.

AC Q61678;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE T-CELL LEUKEMIA, HOMEBOX 1 (311) (FRAGMENT).

GN TLX1 OR 311.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C;

RX MEDLINE; 93281593.

RA DEAR T.N., SANCHEZ-GARCIA I., RABBITTS T.H.;

RT "The HOX11 gene encodes a DNA-binding nuclear transcription factor

RT belonging to a distinct family of homeobox genes.";

RL Proc. Natl. Acad. Sci. U.S.A. 90:4431-4435(1993).

DR EMBL; L08620; AAA28614.1; -

DR PFAM: PF00046; homeobox; 1.
FT NON_TER 1 1
SQ SEQUENCE 27 AA; 2986 MW; 08DA4C92 CRC32;

Query Match 52.9%; Score 9; DB 5; Length 27;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
DB 11 RAAL 14

RESULT 11
Q92079 ID Q92079 PRELIMINARY; PRT; 27 AA.

AC Q92079;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE E2 PROTEIN (FRAGMENT).

OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepatitis C-like viruses.

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98411405.

RA MIZUNO M., HIGUCHI T., KANMATSUSE K., ESUMI M.;
RT "Genetic and serological evidence for multiple instances of
RT unrecognized transmission of hepatitis C virus in hemodialysis
RT units.";
RL J. Clin. Microbiol. 36:2926-2931(1998).

DR EMBL; AB001390; BAA35030.1; -
FT NON_TER 1 1
SQ SEQUENCE 27 AA; 2797 MW; 68BFC607 CRC32;

Query Match 52.9%; Score 9; DB 12; Length 27;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
DB 16 RTSL 19

RESULT 12

Q92063 ID Q92063 PRELIMINARY; PRT; 27 AA.

AC Q92063;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE E2 REGION (FRAGMENT).

OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepatitis C-like viruses.

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98411405.

RA MIZUNO M., HIGUCHI T., KANMATSUSE K., ESUMI M.;
RT "Genetic and serological evidence for multiple instances of
RT unrecognized transmission of hepatitis C virus in hemodialysis
RT units.";
RL J. Clin. Microbiol. 36:2926-2931(1998).

DR EMBL; AB001409; BAA35048.1; -
FT NON_TER 1 1
SQ SEQUENCE 27 AA; 2783 MW; 239ED5B CRC32;

Query Match 52.9%; Score 9; DB 12; Length 27;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
DB 16 RTSL 19

RESULT 13

Q92058 ID Q92058 PRELIMINARY; PRT; 27 AA.

AC Q92058;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE E2 REGION (FRAGMENT).

OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepatitis C-like viruses.

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98411405.

RA MIZUNO M., HIGUCHI T., KANMATSUSE K., ESUMI M.;
RT "Genetic and serological evidence for multiple instances of
RT unrecognized transmission of hepatitis C virus in hemodialysis
RT units.";
RL J. Clin. Microbiol. 36:2926-2931(1998).

DR EMBL; AB001414; BAA35053.1; -
FT NON_TER 1 1
SQ SEQUENCE 27 AA; 2804 MW; 91F8FD5B CRC32;

Query Match 52.9%; Score 9; DB 12; Length 27;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
DB 16 RTSL 19

RESULT 14

Q37005 ID Q37005 PRELIMINARY; PRT; 28 AA.

AC Q37005; O03579; O03588;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE ORF28.

OS Oryza sativa (Rice).

OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
OC Poaceae; Oryza.

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. NIPPONBARE.;

RA SUGIURA M.;
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. NIPPONBARE.;

RX MEDLINE; 99364698.

RA HIRATSUKA J., SHIMADA H., WHITTIER R., ISHIBASHI T., SAKAMOTO M.,
RA MORI M., KONDO C., HONJI Y., SUN C.R., MENG B.Y., LI Y.Q., KANNO A.,

RA NISHIZAWA Y., HIRAI A., SHINOZAKI K., SUGIURA M.;
RT "The complete sequence of the rice (Oryza sativa) chloroplast genome:
RT intermolecular recombination between distinct tRNA genes accounts for
RT a major plastid DNA inversion during the evolution of the cereals.";
RT Mol. Gen. Genet. 217:185-194(1989).

RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE.;
 RX MEDLINE; 91212240.
 RA SHIMADA H., SUGIURA M.;
 RT "Fine structural features of the chloroplast genome: comparison of the
 RL sequenced chloroplast genomes.";
 RN Nucleic Acids Res. 19:983-995(1991).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE.;
 RX MEDLINE; 94073993.
 RA MORTON B.R., CLEGG M.T.;
 RT "A chloroplast DNA mutational hotspot and gene conversion in a
 RL noncoding region near rbcL in the grass family (Poaceae).";
 DR Curr. Genet. 24:357-365(1993).
 DR EMBL; X15901; CAA33922.1; -;
 DR EMBL; X15901; CAA33939.1; -;
 KW Chloroplast.
 SQ SEQUENCE 28 AA; 3419 MW; 66D507A2 CRC32;

Query Match 52.9%; Score 9; DB 8; Length 28;
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 | |
 Db 7 RSSL 10

RESULT 15
 Q00491
 ID Q00491 PRELIMINARY; PRT; 32 AA.
 AC Q00491;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE HYPOTHETICAL 3.6 KD PROTEIN (FRAGMENT).
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M145 SSP. A3(2);
 RX MEDLINE; 9136038.
 RA VAN WEZEL G.P., VIJGENBOOM E., BOSCH L.;
 RT "A comparative study of the ribosomal RNA operons of Streptomyces
 RL coelicolor A3(2) and sequence analysis of rnaA.";
 DR Nucleic Acids Res. 19:4399-4403(1991).
 DR EMBL; X60514; CAA43029.1; -;
 DR PIR; S22304; S22304.
 KW Hypothetical protein.
 FT NON_TER 1
 FT NON_TER 32 32
 SQ SEQUENCE 32 AA; 3617 MW; 05DAB2AB CRC32;

Query Match 52.9%; Score 9; DB 2; Length 32;
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RXXL 4
 | |
 Db 3 RTAL 6

Search completed: February 8, 2000, 13:17:25
 Job time: 32474 sec

THIS PAGE BLANK (USPTO)

Align seq 1/1 to reverse of: A11095 From: 1 to: 15

1 Arg*****Leu 4
|||||
14 CGGACAACTCTC 3

seq_name: gb_pat:AR033507

seq_documentation_block: 15 bp DNA PAT 29-SEP-1999
LOCUS AR033507
DEFINITION Sequence 273 from patent US 5869253.
ACCESSION AR033507
VERSION AR033507.1 GI:5949112
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 15)
AUTHORS Draper,K.G.
TITLE Method and reagent for inhibiting hepatitis C virus replication
JOURNAL Patent: US 5869253-A 273 09-FEB-1999;
FEATURES
Location/Qualifiers
source
1..15
BASE COUNT 1 a 6 c 5 g 3 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x AR033507/rev ..
Align seg 1/1 to reverse of: AR033507 from: 1 to: 15

1 Arg*****Leu 4
|||||
13 CGGGCAACCCCTG 2

seq_name: gb_pat:AR033515

seq_documentation_block: 15 bp DNA PAT 29-SEP-1999
LOCUS AR033515
DEFINITION Sequence 281 from patent US 5869253.
ACCESSION AR033515
VERSION AR033515.1 GI:5949120
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 15)
AUTHORS Draper,K.G.
TITLE Method and reagent for inhibiting hepatitis C virus replication
JOURNAL Patent: US 5869253-A 281 09-FEB-1999;
FEATURES
Location/Qualifiers
source
1..15
BASE COUNT 3 a 3 c 5 g 4 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x AR033515 ..
Align seg 1/1 to: AR033515 from: 1 to: 15

1 Arg*****Leu 4
|||||
4 AGGTCGTCACCTA 15

seq_name: gb_pat:AR033516

seq_documentation_block: 15 bp DNA PAT 29-SEP-1999
LOCUS AR033516
DEFINITION Sequence 282 from patent US 5869253.
ACCESSION AR033516
VERSION AR033516.1 GI:5949121
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 15)
AUTHORS Draper,K.G.
TITLE Method and reagent for inhibiting hepatitis C virus replication
JOURNAL Patent: US 5869253-A 282 09-FEB-1999;
FEATURES
Location/Qualifiers
source
1..15
BASE COUNT 4 a 4 c 4 g 3 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x AR033516 ..
Align seg 1/1 to: AR033516 from: 1 to: 15

1 Arg*****Leu 4
|||||
1 AGGTCGTCACCTA 12

seq_name: gb_pat:AR041315

seq_documentation_block: 15 bp DNA PAT 29-SEP-1999
LOCUS AR041315
DEFINITION Sequence 105 from patent US 5811300.
ACCESSION AR041315
VERSION AR041315.1 GI:5961811
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 15)
AUTHORS Sullivan,S., Draper,K., Kisich,K., Stinchcomb,D.T. and McSwiggen,J.
TITLE TNF- α alpha ribozymes
JOURNAL Patent: US 5811300-A 105 22-SEP-1998;
FEATURES
Location/Qualifiers
source
1..15
BASE COUNT 2 a 3 c 5 g 5 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x AR041315 ..
Align seg 1/1 to: AR041315 from: 1 to: 15

1 Arg*****Leu 4


```

- |||||
4 AGGTCTACTTTG 15

seq_name: gb_pat:AR041316
seq_documentation_block: 15 bp DNA PAT 29-SEP-1999
LOCUS AR041316
DEFINITION Sequence 106 from patent US 5811300.
ACCESSION AR041316
VERSION AR041316.1 GI:5961812
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 15)
AUTHORS Sullivan,S., Draper,K., Kisich,K., Stinchcomb,D.T. and McSwiggen,J.
TITLE TNF- $\alpha$ . ribozymes
JOURNAL Patent: US 5811300-A 106 22-SEP-1998;
FEATURES Location/Qualifiers
source 1..15
/organism="unknown"
BASE COUNT 2 a 3 c 5 g 5 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x AR041316 ..
Align seg 1/1 to: AR041316 from: 1 to: 15

1 Arg*****Leu 4
|||||
2 AGGTCTACTTTG 13

seq_name: gb_pat:AR041841
seq_documentation_block: 15 bp DNA PAT 29-SEP-1999
LOCUS AR041841
DEFINITION Sequence 631 from patent US 5811300.
ACCESSION AR041841
VERSION AR041841.1 GI:5962337
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 15)
AUTHORS Sullivan,S., Draper,K., Kisich,K., Stinchcomb,D.T. and McSwiggen,J.
TITLE TNF- $\alpha$ . ribozymes
JOURNAL Patent: US 5811300-A 631 22-SEP-1998;
FEATURES Location/Qualifiers
source 1..15
/organism="unknown"
BASE COUNT 2 a 3 c 5 g 5 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x AR041841 ..
Align seg 1/1 to: AR041841 from: 1 to: 15

1 Arg*****Leu 4
|||||
2 AGGTCTACTTTG 13

seq_name: gb_pat:AR041316
seq_documentation_block: 15 bp DNA PAT 29-SEP-1999
LOCUS AR041316
DEFINITION Sequence 106 from patent US 5811300.
ACCESSION AR041316
VERSION AR041316.1 GI:5961812
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 15)
AUTHORS Sullivan,S., Draper,K., Kisich,K., Stinchcomb,D.T. and McSwiggen,J.
TITLE TNF- $\alpha$ . ribozymes
JOURNAL Patent: US 5811300-A 106 22-SEP-1998;
FEATURES Location/Qualifiers
source 1..15
/organism="unknown"
BASE COUNT 2 a 3 c 5 g 5 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x AR041316 ..
Align seg 1/1 to: AR041316 from: 1 to: 15

1 Arg*****Leu 4
|||||
2 AGGTCTACTTTG 13

seq_name: gb_pat:AR041841
seq_documentation_block: 15 bp DNA PAT 29-SEP-1999
LOCUS AR041841
DEFINITION Sequence 631 from patent US 5811300.
ACCESSION AR041841
VERSION AR041841.1 GI:5962337
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 15)
AUTHORS Sullivan,S., Draper,K., Kisich,K., Stinchcomb,D.T. and McSwiggen,J.
TITLE TNF- $\alpha$ . ribozymes
JOURNAL Patent: US 5811300-A 631 22-SEP-1998;
FEATURES Location/Qualifiers
source 1..15
/organism="unknown"
BASE COUNT 2 a 3 c 5 g 5 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x AR041841 ..
Align seg 1/1 to: AR041841 from: 1 to: 15

1 Arg*****Leu 4
|||||
2 AGGTCTACTTTG 13

seq_name: gb_pat:AR041842
seq_documentation_block: 15 bp DNA PAT 29-SEP-1999
LOCUS AR041842
DEFINITION Sequence 632 from patent US 5811300.
ACCESSION AR041842
VERSION AR041842.1 GI:5962338
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 15)
AUTHORS Sullivan,S., Draper,K., Kisich,K., Stinchcomb,D.T. and McSwiggen,J.
TITLE TNF- $\alpha$ . ribozymes
JOURNAL Patent: US 5811300-A 632 22-SEP-1998;
FEATURES Location/Qualifiers
source 1..15
/organism="unknown"
BASE COUNT 3 a 3 c 4 g 5 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x AR041842 ..
Align seg 1/1 to: AR041842 from: 1 to: 15

1 Arg*****Leu 4
|||||
2 AGGTCTACTTTG 13

seq_name: gb_pat:AR041843
seq_documentation_block: 15 bp DNA PAT 29-SEP-1999
LOCUS AR041843
DEFINITION Sequence 633 from patent US 5811300.
ACCESSION AR041843
VERSION AR041843.1 GI:5962339
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 15)
AUTHORS Sullivan,S., Draper,K., Kisich,K., Stinchcomb,D.T. and McSwiggen,J.
TITLE TNF- $\alpha$ . ribozymes
JOURNAL Patent: US 5811300-A 633 22-SEP-1998;
FEATURES Location/Qualifiers
source 1..15
/organism="unknown"
BASE COUNT 3 a 3 c 4 g 5 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x AR041843 ..
Align seg 1/1 to: AR041843 from: 1 to: 15

1 Arg*****Leu 4
|||||
2 AGGTCTACTTTG 13
```

```
seq_name: gb_pat:AR041844
seq_documentation_block:
LOCUS AR041844 15 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 634 from patent US 5811300.
ACCESSION AR041844
VERSION AR041844.1 GI:5962340
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 15)
AUTHORS Sullivan,S., Draper,K., Kisch,K., Stinchcomb,D.T. and McSwiggen,J.
TITLE TNF.alpha. ribozymes
JOURNAL Patent: US 5811300-A 634 22-SEP-1998;
FEATURES
Location/Qualifiers
source 1..15
/organism="unknown"
BASE COUNT 3 a 3 c 4 g 5 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x AR041844 ..
Align seg 1/1 to: AR041844 from: 1 to: 15

1 Arg*****Leu 4
|||||
2 AGGTCTACTTG 13

seq_name: gb_pat:I57736
seq_documentation_block:
LOCUS I57736 15 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 273 from patent US 5610054.
ACCESSION I57736
VERSION I57736.1 GI:2482800
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 15)
AUTHORS Draper,K.G.
TITLE Enzymatic RNA molecule targeted against Hepatitis C virus
JOURNAL Patent: US 5610054-A 273 11-MAR-1997;
FEATURES
Location/Qualifiers
source 1..15
/organism="unknown"
BASE COUNT 1 a 6 c 5 g 3 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x I57736/rev ..
Align seg 1/1 to reverse of: I57736 from: 1 to: 15

1 Arg*****Leu 4
|||||
13 CGGCGACCCCTG 2

seq_name: gb_pat:I57744
seq_documentation_block:
LOCUS I57744 15 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 281 from patent US 5610054.
ACCESSION I57744
VERSION I57744.1 GI:2482808
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 15)
AUTHORS Draper,K.G.
TITLE Enzymatic RNA molecule targeted against Hepatitis C virus
JOURNAL Patent: US 5610054-A 281 11-MAR-1997;
FEATURES
Location/Qualifiers
source 1..15
/organism="unknown"
BASE COUNT 3 a 3 c 5 g 4 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x I57744 ..
Align seg 1/1 to: I57744 from: 1 to: 15

1 Arg*****Leu 4
|||||
4 AGGTCGCTACTA 15

seq_name: gb_pat:I57745
seq_documentation_block:
LOCUS I57745 15 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 282 from patent US 5610054.
ACCESSION I57745
VERSION I57745.1 GI:2482809
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 15)
AUTHORS Draper,K.G.
TITLE Enzymatic RNA molecule targeted against Hepatitis C virus
JOURNAL Patent: US 5610054-A 282 11-MAR-1997;
FEATURES
Location/Qualifiers
source 1..15
/organism="unknown"
BASE COUNT 4 a 4 c 4 g 3 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x I57745 ..
Align seg 1/1 to: I57745 from: 1 to: 15

1 Arg*****Leu 4
|||||
1 AGGTCGCTACTA 12

seq_name: gb_pat:A42606
```

seq_documentation_block:
LOCUS A42606 16 bp DNA PAT 06-MAR-1997
DEFINITION Sequence 124 from Patent WO9502051.
ACCESSION A42606
VERSION A42606.1 GI:2298055
KEYWORDS
SOURCE
ORGANISM
unidentified.
unclassified.
REFERENCE 1 (bases 1 to 16)
AUTHORS Schlingensiepen,G., Schlingensiepen,R., Schlingensiepen,K. and
Brysch,W.
TITLE A PHARMACEUTICAL COMPOSITION COMPRISING ANTISENSE-NUCLEIC ACID FOR
PREVENTION AND/OR TREATMENT OF NEURONAL INJURY, DEGENERATION AND
CELL DEATH AND FOR THE TREATMENT OF NEOPLASMS
JOURNAL Patent: WO 9502051-A 124 19-JAN-1995;
COMMENT BIOGNOSTIK GES FUEER BIOMOLEKUL (DE)
FEATURES Other publication AU 7345694 950206.
Location/Qualifiers
source
1..16
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 4 a 2 c 5 g 5 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x A42606/rev ..
Align seg 1/1 to reverse of: A42606 from: 1 to: 16

1 Arg*****Leu 4
|||||||
12 AGAGTACTCTC 1

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-1 to: N_Geneseq_36:* out_format : pfs
 Date: Feb 8, 2000 1:27 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+ p2n.model -DEV=klp
 -Q=/cgn1_1/USPTO_spool/US08653294/runat_04022000_160701_15807/app_query.fasta.1
 -DB=N_Geneseq_36 -OFMT=fastap -SUFFIX=ring -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT
 -THREADS=1

Search information block:
 Query: US-08-653-294-1
 Query length: 10
 Database: N_Geneseq_36:*
 Database sequences: 311585
 Database length: 125096042
 Search time (sec): 590.520000

| Sequence | Score | Strid | Orig | ZScore | Escore | Len | Documentation |
|----------------------|-------|-------|-------|--------|--------|-----|-----------------------------------|
| N_Geneseq_36:V40924 | - | 9.00 | 98.28 | 705.67 | 12 | 1 | Primer AML1EVI:4905L12 for abno |
| N_Geneseq_36:V65428 | - | 9.00 | 98.28 | 705.67 | 12 | 1 | Primer PBS7-9C used in the cours |
| N_Geneseq_36:V65424 | - | 9.00 | 98.28 | 705.67 | 12 | 1 | Primer PBS7-8B used in the cours |
| N_Geneseq_36:V65478 | - | 9.00 | 98.28 | 705.67 | 12 | 1 | Primer PBS800-89E used in the cc |
| N_Geneseq_36:V65479 | - | 9.00 | 98.28 | 705.67 | 12 | 1 | Primer PBS800-89F used in the cc |
| N_Geneseq_36:V65486 | - | 9.00 | 98.28 | 705.67 | 12 | 1 | Primer PBS800-910D used in the c |
| N_Geneseq_36:V65487 | - | 9.00 | 98.28 | 705.67 | 12 | 1 | Primer PBS800-910E used in the c |
| N_Geneseq_36:V65488 | - | 9.00 | 98.28 | 705.67 | 12 | 1 | Primer PBS800-910F used in the c |
| N_Geneseq_36:V674335 | - | 9.00 | 97.00 | 831.22 | 14 | 1 | Antisense oligonucleotide #17 cc |
| N_Geneseq_36:V67435 | - | 9.00 | 97.00 | 831.22 | 14 | 1 | Invader oligonucleotide IT-1. TH |
| N_Geneseq_36:V674719 | - | 9.00 | 97.00 | 831.22 | 14 | 1 | Invader oligonucleotide IT-1A4. |
| N_Geneseq_36:V38651 | - | 9.00 | 97.00 | 831.22 | 14 | 1 | Human ICAM-1, E-selectin, VCAM-1 |
| N_Geneseq_36:V48937 | - | 9.00 | 97.00 | 831.22 | 14 | 1 | TGF-beta2 antisense oligonucleot |
| N_Geneseq_36:V48608 | - | 9.00 | 97.00 | 831.22 | 14 | 1 | jund gene antisense oligonucleot |
| N_Geneseq_36:V48594 | - | 9.00 | 97.00 | 831.22 | 14 | 1 | jund gene antisense oligonucleot |
| N_Geneseq_36:V53960 | - | 9.00 | 97.00 | 831.22 | 14 | 1 | Nucleotide sequence of the invad |
| N_Geneseq_36:V53961 | - | 9.00 | 97.00 | 831.22 | 14 | 1 | Nucleotide sequence of the olig |
| N_Geneseq_36:V65877 | - | 9.00 | 97.00 | 831.22 | 14 | 1 | Invader oligonucleotide IT-1. De |
| N_Geneseq_36:V65878 | - | 9.00 | 97.00 | 831.22 | 14 | 1 | Invader oligonucleotide IT-1A4. |
| N_Geneseq_36:V65879 | - | 9.00 | 96.43 | 894.42 | 15 | 1 | Antisense oligomer C3 [1588] whic |
| N_Geneseq_36:V65880 | - | 9.00 | 96.43 | 894.42 | 15 | 1 | Human CERP HH ribozyme target s |
| N_Geneseq_36:V65881 | - | 9.00 | 96.43 | 894.42 | 15 | 1 | Rabbit CERP HH ribozyme target s |
| N_Geneseq_36:V65882 | - | 9.00 | 96.43 | 894.42 | 15 | 1 | Rabbit CERP HH ribozyme target s |
| N_Geneseq_36:V65883 | - | 9.00 | 96.43 | 894.42 | 15 | 1 | Rabbit CERP HH ribozyme target s |
| N_Geneseq_36:V65884 | - | 9.00 | 96.43 | 894.42 | 15 | 1 | Human TNF-alpha hammerhead riboz |
| N_Geneseq_36:V65885 | - | 9.00 | 96.43 | 894.42 | 15 | 1 | Human TNF-alpha hammerhead riboz |
| N_Geneseq_36:V65886 | - | 9.00 | 96.43 | 894.42 | 15 | 1 | Mouse TNF-a hammerhead ribozyme |
| N_Geneseq_36:V65887 | - | 9.00 | 96.43 | 894.42 | 15 | 1 | Mouse TNF-a hammerhead ribozyme |
| N_Geneseq_36:V65888 | - | 9.00 | 96.43 | 894.42 | 15 | 1 | Target sequence with sequence hc |
| N_Geneseq_36:V65889 | - | 9.00 | 96.43 | 894.42 | 15 | 1 | Sequencing primer SSK-07 for S.e |
| N_Geneseq_36:V65890 | - | 9.00 | 95.89 | 957.89 | 16 | 1 | Primer based on plasmid construc |
| N_Geneseq_36:V65891 | - | 9.00 | 95.89 | 957.89 | 16 | 1 | c-fos antisense oligonucleotide. |
| N_Geneseq_36:V65892 | - | 9.00 | 95.89 | 957.89 | 16 | 1 | Primer A (Group 13, set B) for m |
| N_Geneseq_36:V65893 | - | 9.00 | 95.89 | 957.89 | 16 | 1 | G. oxydans T100 L-sorbose dehyd |
| N_Geneseq_36:V65894 | - | 9.00 | 95.89 | 957.89 | 16 | 1 | Bacteriophage T4 permuted gene 3 |
| N_Geneseq_36:V65895 | - | 9.00 | 95.89 | 957.89 | 16 | 1 | Adaptor directed primer Hnd.pr.5 |
| N_Geneseq_36:V65896 | - | 9.00 | 95.89 | 957.89 | 16 | 1 | Adaptor directed primer Hnd.pr.6 |
| N_Geneseq_36:V65897 | - | 9.00 | 95.89 | 957.89 | 16 | 1 | HIV-1 gag binding oligonucleotid |
| N_Geneseq_36:V65898 | - | 9.00 | 95.89 | 957.89 | 16 | 1 | Hepatitis C virus recognition se |
| N_Geneseq_36:V65899 | - | 9.00 | 95.89 | 957.89 | 16 | 1 | HIV-1 gag gene antisense oligonu |

N_Geneseq_36:V30249 + 9.00 95.89 957.89 16 ! Labelled oligonucleotide used
 N_Geneseq_36:V49885 - 9.00 95.89 957.89 16 ! SV40 promoter Gal4 DNA bindin
 N_Geneseq_36:V45767 - 9.00 95.89 957.89 16 ! Capture probe 13. Reaction su
 N_Geneseq_36:V02670 - 9.00 95.89 957.89 16 ! BP-893493 Seq ID 4. New endot

seq_name: N_Geneseq_36:V40924

seq_documentation_block:

ID V40924 standard; DNA; 12 BP.
 AC V40924;
 DT 25-SEP-1998 (first entry)
 DE Primer AML1EVI:4905L12 for abnormality detection.
 KW PCR primer; chromosomal abnormality; abnormality detection; leukaemia;
 KW lymphoma; carcinoma; adenocarcinoma; sarcoma; glioma; neuroblastoma;
 KW medullablastoma; malignant melanoma; malignant neoplastic condition; ss.
 OS Synthetic.
 PS Homo sapiens.
 PN W09824928-A2.
 PD 11-JUN-1998.
 PF 08-DEC-1997; DR0556.
 PR 06-DEC-1996; DK-001401.
 PA (PALL/) PALLISGAARD N.
 PI Hokland P, Pallisgaard N;
 DR WPI: 98-333344/29
 PT Detection of chromosomal abnormalities - by subjecting patient
 PT sample nucleic acids to a multiplex molecular amplification
 PT procedure using primers specific for characteristic nucleic acid
 PT sequence
 PS Claim 73; Page 66; 126pp; English.
 CC This sequence represents a primer used in the method of the invention for
 CC the detection of the presence or absence of chromosomal abnormalities,
 CC each abnormality being associated with a condition in a subject and each
 CC being defined by at least one characteristic nucleic acid sequence. The
 CC method comprises: (a) obtaining a sample of nucleic acids derived from a
 CC subject which may harbour one of the chromosomal abnormalities;
 CC (b) subjecting the sample to a multiplex molecular amplification (MMA)
 CC procedure, where a number of the characteristic sequences, if present in
 CC a sufficient amount, will be amplified; (c) retrieving the product(s)
 CC from step (b), and detecting the presence and/or absence of an amplicon
 CC characteristic of the abnormal sequences to detect the presence or
 CC absence of corresponding chromosomal abnormalities; where the MMA
 CC procedure comprises the use of at least 7 mutually distinct primers (MDP)
 CC in one single reaction mixture, each of the primers defining an end of at
 CC least one characteristic nucleic acid sequence, and where at least one of
 CC the primers defines the first end of at least two characteristic nucleic
 CC acid sequences, the characteristic nucleic acid sequences each being
 CC determined in their opposite ends by MDP selected from the remainder of
 CC the MDP. The methods can be used for detecting chromosomal abnormalities
 CC associated with diseases including numerous leukaemia's, lymphoma's,
 CC carcinoma's, adenocarcinoma's, sarcoma's, glioma's, neuroblastoma's,
 CC medullablastoma, malignant melanoma, and malignant neoplastic conditions.
 SQ Sequence 12 BP; 2 A; 3 C; 4 T;

alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-1 x V40924/rev
 Align seg 1/1 to reverse of: V40924 from: 1 to: 12

1 Arg*****Leu 4
 |||||
 12 AGAGCAGCCTTA 1

seq_name: N_Geneseq_36:V65428

seq_documentation_block:

ID V65428 standard; DNA; 12 BP.
 AC V65428;

DT 08-DEC-1998 (first entry)
 DE Primer pBS7-9C used in the course of the invention.
 KW Nucleic acid determination; hybridisation; probe; mismatch; SBH;
 KW sequencing by hybridisation; PCR primer; ss.
 OS Synthetic.
 PN J10243785-A.
 PD 14-SEP-1998.
 PF 03-MAR-1997; 047821.
 PR 03-MAR-1997; JP-047821.
 PA (BUNS-) BUNSHI BIOHOTOONICS KENKYUSHO KK.
 DR WPI: 98-549781/47.
 PT Determination of nucleic acid base sequence - is sensitive and rapid
 PT without mismatch in hybridisation as in sequencing by hybridisation
 PT method
 PS Examples; Page 8; 20pp; Japanese.
 CC Sequences shown in V65401 to V65580 represent PCR primers used in the
 CC course of the invention which provides a method for determining a single
 CC stranded nucleic acid base sequence. The method comprises separation of
 CC 4k oligonucleotide probe as a primer from all combinations of k base
 CC sequences and hybridising the probe and the nucleic acid to be tested.
 CC The probe is elongated to make a primer using the nucleic acid to be
 CC tested as a template and the elongated primer is determined. The base
 CC sequence of the nucleic acid is determined based on the elongated
 CC amount. The method allows sensitive and rapid determination of nucleic
 CC acid base sequence without mismatch in hybridisation as in sequencing
 CC by hybridisation (SBH) method.
 SQ Sequence 12 BP; 2 A; 3 C; 5 G; 2 T;

 alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

 alignment_block:
 US-08-653-294-1 x V65428/rev ..
 Align seg 1/1 to reverse of: V65428 from: 1 to: 12

 1 Arg*****Leu 4
 |||||
 12 CGCAGCCGCTTA 1

 seq_name: N_Geneseq_36:V65424

 seq_documentation_block:
 ID V65424 standard; DNA; 12 BP.
 AC V65424;
 DT 08-DEC-1998 (first entry)
 DE Primer pBS7-8B used in the course of the invention.
 KW Nucleic acid determination; hybridisation; probe; mismatch; SBH;
 KW sequencing by hybridisation; PCR primer; ss.
 OS Synthetic.
 PN J10243785-A.
 PD 14-SEP-1998.
 PF 03-MAR-1997; 047821.
 PR 03-MAR-1997; JP-047821.
 PA (BUNS-) BUNSHI BIOHOTOONICS KENKYUSHO KK.
 DR WPI: 98-549781/47.
 PT Determination of nucleic acid base sequence - is sensitive and rapid
 PT without mismatch in hybridisation as in sequencing by hybridisation
 PT method
 PS Examples; Page 8; 20pp; Japanese.
 CC Sequences shown in V65401 to V65580 represent PCR primers used in the
 CC course of the invention which provides a method for determining a single
 CC stranded nucleic acid base sequence. The method comprises separation of
 CC 4k oligonucleotide probe as a primer from all combinations of k base
 CC sequences and hybridising the probe and the nucleic acid to be tested.
 CC The probe is elongated to make a primer using the nucleic acid to be
 CC tested as a template and the elongated primer is determined. The base
 CC sequence of the nucleic acid is determined based on the elongated
 CC amount. The method allows sensitive and rapid determination of nucleic
 CC acid base sequence without mismatch in hybridisation as in sequencing

CC by hybridisation (SBH) method.
 SQ Sequence 12 BP; 3 A; 2 C; 6 G; 1 T;

 alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

 alignment_block:
 US-08-653-294-1 x V65424/rev ..
 Align seg 1/1 to reverse of: V65424 from: 1 to: 12

 1 Arg*****Leu 4
 |||||
 12 CGTCGCCCTTA 1

 seq_name: N_Geneseq_36:V65478

 seq_documentation_block:
 ID V65478 standard; DNA; 12 BP.
 AC V65478;
 DT 08-DEC-1998 (first entry)
 DE Primer pBS800-89E used in the course of the invention.
 KW Nucleic acid determination; hybridisation; probe; mismatch; SBH;
 KW sequencing by hybridisation; PCR primer; ss.
 OS Synthetic.
 PN J10243785-A.
 PD 14-SEP-1998.
 PF 03-MAR-1997; 047821.
 PR 03-MAR-1997; JP-047821.
 PA (BUNS-) BUNSHI BIOHOTOONICS KENKYUSHO KK.
 DR WPI: 98-549781/47.
 PT Determination of nucleic acid base sequence - is sensitive and rapid
 PT without mismatch in hybridisation as in sequencing by hybridisation
 PT method
 PS Examples; Page 10; 20pp; Japanese.
 CC Sequences shown in V65401 to V65580 represent PCR primers used in the
 CC course of the invention which provides a method for determining a single
 CC stranded nucleic acid base sequence. The method comprises separation of
 CC 4k oligonucleotide probe as a primer from all combinations of k base
 CC sequences and hybridising the probe and the nucleic acid to be tested.
 CC The probe is elongated to make a primer using the nucleic acid to be
 CC tested as a template and the elongated primer is determined. The base
 CC sequence of the nucleic acid is determined based on the elongated
 CC amount. The method allows sensitive and rapid determination of nucleic
 CC acid base sequence without mismatch in hybridisation as in sequencing
 CC by hybridisation (SBH) method.
 SQ Sequence 12 BP; 2 A; 2 C; 6 G; 2 T;

 alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

 alignment_block:
 US-08-653-294-1 x V65478/rev ..
 Align seg 1/1 to reverse of: V65478 from: 1 to: 12

 1 Arg*****Leu 4
 |||||
 12 CGCAGCCGCTTA 1

 seq_name: N_Geneseq_36:V65479

 seq_documentation_block:
 ID V65479 standard; DNA; 12 BP.
 AC V65479;
 DT 08-DEC-1998 (first entry)
 DE Primer pBS800-89F used in the course of the invention.

```
KW Nucleic acid determination; hybridisation; probe: mismatch; SBH;
KW sequencing by hybridisation; PCR primer; ss.
PN J10243785-A.
PD 14-SEP-1998.
PF 03-MAR-1997; 047821.
PR 03-MAR-1997; JP-047821.
PA (BUNS-) BUNSHI BIOTRONICS KENYUSHO KK.
DR WPI; 98-549781/47.
PT Determination of nucleic acid base sequence - is sensitive and rapid
PT without mismatch in hybridisation as in sequencing by hybridisation
PT method
PS Examples; Page 10; 20pp; Japanese.
CC Sequences shown in V65401 to V65580 represent PCR primers used in the
CC course of the invention which provides a method for determining a single
CC stranded nucleic acid base sequence. The method comprises separation of
CC 4k oligonucleotide probe as a primer from all combinations of k base
CC sequences and hybridising the probe and the nucleic acid to be tested.
CC The probe is elongated to make a primer using the nucleic acid to be
CC tested as a template and the elongated primer is determined. The base
CC sequence of the nucleic acid is determined based on the elongated
CC amount. The method allows sensitive and rapid determination of nucleic
CC acid base sequence without mismatch in hybridisation as in sequencing
CC by hybridisation (SBH) method.
SQ Sequence 12 BP; 2 A; 3 C; 6 G; 1 T;

alignment_scores:
  Quality: 9.00      Length: 4
  Ratio: 2.250      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x V65479/rev ..
Align seg 1/1 to reverse of: V65479 from: 1 to: 12

1 Arg*****Leu 4
|||||
12 CGCGCGCGCTTA 1

seq_name: N_Geneseq_36:V65486
seq_documentation_block:
ID V65486 standard; DNA; 12 BP.
AC V65486;
DT 08-DEC-1998 (first entry)
DE Primer PBS800-910D used in the course of the invention.
KW Nucleic acid determination; hybridisation; probe: mismatch; SBH;
KW sequencing by hybridisation; PCR primer; ss.
OS Synthetic.
PN J10243785-A.
PD 14-SEP-1998.
PF 03-MAR-1997; 047821.
PR 03-MAR-1997; JP-047821.
PA (BUNS-) BUNSHI BIOTRONICS KENYUSHO KK.
DR WPI; 98-549781/47.
PT Determination of nucleic acid base sequence - is sensitive and rapid
PT without mismatch in hybridisation as in sequencing by hybridisation
PT method
PS Examples; Page 10; 20pp; Japanese.
CC Sequences shown in V65401 to V65580 represent PCR primers used in the
CC course of the invention which provides a method for determining a single
CC stranded nucleic acid base sequence. The method comprises separation of
CC 4k oligonucleotide probe as a primer from all combinations of k base
CC sequences and hybridising the probe and the nucleic acid to be tested.
CC The probe is elongated to make a primer using the nucleic acid to be
CC tested as a template and the elongated primer is determined. The base
CC sequence of the nucleic acid is determined based on the elongated
CC amount. The method allows sensitive and rapid determination of nucleic
CC acid base sequence without mismatch in hybridisation as in sequencing
CC by hybridisation (SBH) method.
SQ Sequence 12 BP; 2 A; 3 C; 4 G; 3 T;

alignment_scores:
  Quality: 9.00      Length: 4
  Ratio: 2.250      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x V65479/rev ..
Align seg 1/1 to reverse of: V65479 from: 1 to: 12

1 Arg*****Leu 4
|||||
12 CGCGCGCGCTTA 1

seq_name: N_Geneseq_36:V65486
seq_documentation_block:
ID V65486 standard; DNA; 12 BP.
AC V65486;
DT 08-DEC-1998 (first entry)
DE Primer PBS800-910D used in the course of the invention.
KW Nucleic acid determination; hybridisation; probe: mismatch; SBH;
KW sequencing by hybridisation; PCR primer; ss.
OS Synthetic.
PN J10243785-A.
PD 14-SEP-1998.
PF 03-MAR-1997; 047821.
PR 03-MAR-1997; JP-047821.
PA (BUNS-) BUNSHI BIOTRONICS KENYUSHO KK.
DR WPI; 98-549781/47.
PT Determination of nucleic acid base sequence - is sensitive and rapid
PT without mismatch in hybridisation as in sequencing by hybridisation
PT method
PS Examples; Page 10; 20pp; Japanese.
CC Sequences shown in V65401 to V65580 represent PCR primers used in the
CC course of the invention which provides a method for determining a single
CC stranded nucleic acid base sequence. The method comprises separation of
CC 4k oligonucleotide probe as a primer from all combinations of k base
CC sequences and hybridising the probe and the nucleic acid to be tested.
CC The probe is elongated to make a primer using the nucleic acid to be
CC tested as a template and the elongated primer is determined. The base
CC sequence of the nucleic acid is determined based on the elongated
CC amount. The method allows sensitive and rapid determination of nucleic
CC acid base sequence without mismatch in hybridisation as in sequencing
CC by hybridisation (SBH) method.
SQ Sequence 12 BP; 2 A; 3 C; 4 G; 3 T;
```

```
alignment_scores:
  Quality: 9.00      Length: 4
  Ratio: 2.250      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x V65486/rev ..
Align seg 1/1 to reverse of: V65486 from: 1 to: 12

1 Arg*****Leu 4
|||||
12 CGAGCGCGCTTA 1

seq_name: N_Geneseq_36:V65487
seq_documentation_block:
ID V65487 standard; DNA; 12 BP.
AC V65487;
DT 08-DEC-1998 (first entry)
DE Primer PBS800-910E used in the course of the invention.
KW Nucleic acid determination; hybridisation; probe: mismatch; SBH;
KW sequencing by hybridisation; PCR primer; ss.
OS Synthetic.
PN J10243785-A.
PD 14-SEP-1998.
PF 03-MAR-1997; 047821.
PR 03-MAR-1997; JP-047821.
PA (BUNS-) BUNSHI BIOTRONICS KENYUSHO KK.
DR WPI; 98-549781/47.
PT Determination of nucleic acid base sequence - is sensitive and rapid
PT without mismatch in hybridisation as in sequencing by hybridisation
PT method
PS Examples; Page 10; 20pp; Japanese.
CC Sequences shown in V65401 to V65580 represent PCR primers used in the
CC course of the invention which provides a method for determining a single
CC stranded nucleic acid base sequence. The method comprises separation of
CC 4k oligonucleotide probe as a primer from all combinations of k base
CC sequences and hybridising the probe and the nucleic acid to be tested.
CC The probe is elongated to make a primer using the nucleic acid to be
CC tested as a template and the elongated primer is determined. The base
CC sequence of the nucleic acid is determined based on the elongated
CC amount. The method allows sensitive and rapid determination of nucleic
CC acid base sequence without mismatch in hybridisation as in sequencing
CC by hybridisation (SBH) method.
SQ Sequence 12 BP; 2 A; 4 C; 4 G; 2 T;

alignment_scores:
  Quality: 9.00      Length: 4
  Ratio: 2.250      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x V65487/rev ..
Align seg 1/1 to reverse of: V65487 from: 1 to: 12

1 Arg*****Leu 4
|||||
12 CGGAGCGCGCTTA 1

seq_name: N_Geneseq_36:V65488
seq_documentation_block:
ID V65488 standard; DNA; 12 BP.
AC V65488;
DT 08-DEC-1998 (first entry)
DE Primer PBS800-910F used in the course of the invention.
KW Nucleic acid determination; hybridisation; probe: mismatch; SBH;
KW sequencing by hybridisation; PCR primer; ss.
```

OS Synthetic.
 PN J10243785-A.
 PD 14-SEP-1998.
 PE 03-MAR-1997; 047821.
 PR 03-MAR-1997; JP-047821.
 PA (BUNSHI) BUNSHI BIOTONICS KENKYUSHO KK.
 DR WPI; 98-549781/47.
 PT Determination of nucleic acid base sequence - is sensitive and rapid
 PT without mismatch in hybridisation as in sequencing by hybridisation
 PT method

PS Examples: Page 10: 20pp: Japanese.
 CC Sequences shown in V65401 to V65580 represent PCR primers used in the
 CC course of the invention which provides a method for determining a single
 CC stranded nucleic acid base sequence. The method comprises separation of
 CC 4k oligonucleotide probe as a primer from all combinations of k base
 CC sequences and hybridising the probe and the nucleic acid to be tested.
 CC The probe is elongated to make a primer using the nucleic acid to be
 CC tested as a template and the elongated primer is determined. The base
 CC sequence of the nucleic acid is determined based on the elongated
 CC amount. The method allows sensitive and rapid determination of nucleic
 CC acid base sequence without mismatch in hybridisation as in sequencing
 CC by hybridisation (SBH) method.
 SQ Sequence 12 BP; 3 A; 3 C; 4 G; 2 T;

alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-08-653-294-1 x V65488/rev ..

Align seg 1/1 to reverse of: V65488 from: 1 to: 12

1 Arg*****Leu 4
 |||||
 12 CGTAGCGCCTTA 1

seq_name: N_Geneseq_36:T64335

seq_documentation_block:
 ID T64335 standard; DNA; 14 BP.
 AC T64335;
 DT 21-MAY-1997 (first entry)
 DE Antisense oligonucleotide #17 complementary to human VCAM-1 mRNA.
 KW Human; vascular cell adhesion molecule; VCAM-1; antisense;
 KW septic shock; downregulation; inflammation; leukocyte adhesion; ss.
 OS Synthetic.
 PN US5596090-A.
 PD 21-JAN-1997.
 PF 24-JUL-1992; 918256.
 PR 24-JUL-1992; US-918256.
 PR 12-OCT-1993; US-137701.
 PA (USNA) US SEC OF NAVY.
 PI Bradley MO, Hoke GB, Lee C, Williams TJ;
 DR WPI; 97-107618/10.
 PT Antisense oligonucleotide(s) for treating septic shock - with
 PT sequence complementary to VCAM-1 mRNA transcript
 PS Claim 1: Column 29; 18pp: English.
 CC The present sequence is that of an antisense oligonucleotide
 CC complementary to a region in the precursor or mature mRNA of human
 CC vascular cell adhesion molecule VCAM-1. The antisense oligonucleotide
 CC (preferably containing phosphorothioate linkages) is used for
 CC downregulating VCAM-1 synthesis which in turn results in a reduction
 CC in adhesion of leukocytes to the endothelium and hence to a reduced
 CC inflammatory response.
 SQ Sequence 14 BP; 2 A; 5 C; 6 G; 1 T;

alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-08-653-294-1 x T64335/rev ..

Align seg 1/1 to reverse of: T64335 from: 1 to: 14

1 Arg*****Leu 4
 |||||
 14 CGGCGCTCACTG 3

seq_name: N_Geneseq_36:T76718

seq_documentation_block:
 ID T76718 standard; DNA; 14 BP.
 AC T76718;
 DT 14-APR-1998 (first entry)
 DE Invader oligonucleotide IT-1.
 KW Nucleic acid cleavage; DNA cleavage; RNA cleavage; nuclease;
 KW invader directed cleavage; cleavage A/G; ss.
 OS Synthetic.
 PN WO9727214-A1.
 PD 31-JUL-1997.
 PF 22-JAN-1997; U01072.
 PR 02-DEC-1996; US-759038.
 PR 24-JAN-1996; US-599491.
 PR 12-JUL-1996; US-682853.
 PR 29-NOV-1996; US-756386.
 PR 02-DEC-1996; US-758314.
 PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
 PI Brow MAD, Dahlberg JE, Hall JG, Kaiser MW, Lyamichev VI,
 PI Olive DM, Prudent JR;
 DR WPI; 97-393613/36.
 PT Thermostable structure-specific nuclease(s) - used for detection and
 PT characterisation of nucleic acid sequences and variations in nucleic
 PT acid sequences
 PS Example 37: Page 301; 457pp: English.
 CC Cleavage of test oligonucleotide IT-2 (T76717) was examined in the
 CC presence of Invader (TM) oligonucleotides IT-1 (T76718) and IT-1A4
 CC (T76719). IT-1 is fully complementary to the 3' arm of IT-2,
 CC whereas IT-1A4 has a T to A substitution that results in an A/A
 CC mismatch in the Invader-target duplex. Experiments were performed
 CC to show that the presence of a mismatch in the Invader
 CC oligonucleotide decreases the cleavage activity of cleavage A/G
 CC nuclease (see T76643). The invention relates to means for the
 CC detection and characterisation of nucleic acid (NA) sequences
 CC and variations in NA sequences. It also relates to methods for
 CC forming a NA cleavage structure on a target sequence and cleaving
 CC the NA cleavage structure in a site-specific manner. The 5'
 CC nuclease activity of various enzymes (see W24210-13) is used to
 CC cleave the target-dependent cleavage structure, thereby indicating
 CC the presence of specific NA sequences or specific variations of
 CC them. The invader directed cleavage reaction of the invention is
 CC an ideal direct detection method that combines the advantages of
 CC direct detection assays with the specificity of a dual or
 CC tri-oligonucleotide hybridisation assay.
 SQ Sequence 14 BP; 3 A; 3 C; 4 G; 4 T;

alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-08-653-294-1 x T76718 ..

Align seg 1/1 to: T76718 from: 1 to: 14

1 Arg*****Leu 4
 |||||
 3 CGAGCGCTTTG 14


```

seq_name: N_Geneseq_36:T76719
seq_documentation_block:
ID T76719 standard; DNA; 14 BP.
AC T76719;
DT 14-APR-1998 (first entry)
DE Invader oligonucleotide IT-1A4.
KW Nucleic acid cleavage; DNA cleavage; RNA cleavage; nuclease;
KW invader directed cleavage; cleavage A/G; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT mutation 11
FT /tag= a
FT /note= "T to A substitution"
PN W09727214-A1.
PD 31-JUL-1997.
PF 22-JAN-1997; U01072.
PR 02-DEC-1996; US-759038.
PR 24-JAN-1996; US-599491.
PR 12-JUL-1996; US-682853.
PR 29-NOV-1996; US-756386.
PR 02-DEC-1996; US-758314.
PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
PI Brow MAD, Dahlberg JE, Hall JG, Kaiser MW, Lyamichiev VI,
PI Olive DM, Prudent JR;
DR WPI: 97-393613/36.
PT Thermostable structure-specific nuclease(s) - used for detection and
PT characterisation of nucleic acid sequences and variations in nucleic
PT acid sequences
PS Example 37: Page 301; 457pp; English.
CC Cleavage of test oligonucleotide IT-2 (T76717) was examined in the
CC presence of invader (TM) oligonucleotides IT-1 (T76718) and IT-1A4
CC (T76719). IT-1 is fully complementary to the 3' arm of IT-2,
CC whereas IT-1A4 has a T to A substitution that results in an A/A
CC mismatch in the invader-target duplex. Experiments were performed
CC to show that the presence of a mismatch in the invader
CC oligonucleotide decreases the cleavage activity of cleavage A/G
CC nuclease (see T76643). The invention relates to means for the
CC detection and characterisation of nucleic acid (NA) sequences
CC and variations in NA sequences. It also relates to methods for
CC forming a NA cleavage structure on a target sequence and cleaving
CC the NA cleavage structure in a site-specific manner. The 5'
CC nuclease activity of various enzymes (see #24210-13) is used to
CC cleave the target-dependent cleavage structure, thereby indicating
CC the presence of specific NA sequences or specific variations of
CC them. The invader directed cleavage reaction of the invention is
CC an ideal direct detection method that combines the advantages of
CC direct detection assays with the specificity of a dual or
CC tri-oligonucleotide hybridisation assay.
SQ Sequence 14 BP; 4 A; 3 C; 4 G; 3 T;

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x T76719 ..
Align seg 1/1 to: T76719 from: 1 to: 14

1 Arg*****Leu 4
|||||
3 CGAGGCTCATTG 14

seq_name: N_Geneseq_36:V38651
seq_documentation_block:
ID V38651 standard; DNA; 14 BP.
AC V38651;
DT 13-OCT-1998 (first entry)
DE Human ICAM-1, E-selectin, VCAM-1 antisense oligonucleotide.
KW ICAM-1; intracellular adhesion molecule-; E-selectin; VCAM-1;
KW vascular cell adhesion molecule-1; antisense; inflammatory;
KW disease; treatment; septic shock; psoriasis; wounds; burns; acne;
KW arthritis; organ rejection; inhibition; expression; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT mutation 11
FT /tag= a
FT /note= "T to A substitution"
PN W09824797-A1.
PD 11-JUN-1998.
PF 02-DEC-1996; U19194.
PR 02-DEC-1996; WO-U19194.
PR (DYAD-) DYAD PHARM CORP.
PI Bradley MO, Hoke GD, Lee C, Williams TJ;
DR WPI: 98-333253/29.
PT Antisense oligonucleotides to ICAM-1, E-selectin or VCAM-1 - useful
PT for treating diseases having an inflammatory component, e.g.
PT psoriasis, wounds and septic shock
PS Claim 8; Page 41; 48pp; English.
CC The sequence is that of an antisense oligonucleotide which is
CC substantially complementary to at least a portion of the pre-
CC or mature RNA transcript of human intracellular adhesion molecule
CC (ICAM), E-selectin or vascular cell adhesion molecule (VCAM).
CC It can be used to inhibit expression of these proteins. Inhibition
CC of these proteins forms the basis for treatment of conditions and
CC diseases that have an inflammatory component, e.g. acne, psoriasis,
CC arthritis, organ rejection, wounds, burns, septic shock or
CC inflammatory complications of septic shock.
SQ Sequence 14 BP; 2 A; 5 C; 6 G; 1 T;

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x V38651/rev ..
Align seg 1/1 to reverse of: V38651 from: 1 to: 14

1 Arg*****Leu 4
|||||
14 CGGGCCTCACTG 3

seq_name: N_Geneseq_36:V48937
seq_documentation_block:
ID V48937 standard; DNA; 14 BP.
AC V48937;
DT 15-OCT-1998 (first entry)
DE TGF-beta2 antisense oligonucleotide TGF-beta2-8.
KW Transforming growth factor-beta2; TGF-beta2;
KW antisense oligonucleotide; modulate; gene expression; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT mutation 11
FT /tag= a
FT /note= "T to A substitution"
PN EP-856579-A1.
PD 05-AUG-1998.
PF 31-JAN-1997; 101531.
PR 31-JAN-1997; EP-101531.
PA (BIOG-) BIOGNOSTIK GES BIOMOLEKULARE DIAGNOSTIK.
PI Brysch W, Schlingensiepen K;
DR WPI: 98-400910/35.
PT Preparation of antisense oligo:nucleotide(s) which lack long runs of
PT consecutive guanosine or inosine - and have specific ratio of
PT residues able to form two or three hydrogen bonds, have greater
PT activity and reduced toxicity, used therapeutically or to modulate
PT growth of cells in culture
PS Claim 10; Fig 8a; 286pp; English.
CC V48930-49007 represent antisense oligonucleotides directed against
CC transforming growth factor-beta2 (TGF-beta2). Of these, only
CC oligonucleotides V48930-67 resulted in significant reduction in
CC TGF-beta 2 protein expression, while oligonucleotides V48968-49007 had
CC little effect. The oligonucleotides exemplify the invention. The
CC specification describes oligonucleotides that contain 8-30 nucleotides,

```

CC which contain at most 8 nucleotides that can each form three hydrogen
 CC bonds to cytosine; do not contain four consecutive nucleotides able to
 CC form three H-bonds each to four consecutive cytosines; do not contain
 CC two sequences of three consecutive nucleotides each able to form three
 CC H-bonds to three consecutive cytosines, and the ratio between residues
 CC able to form two H-bonds each (2R) or three such bonds (3R) is given
 CC by 2R/3R = 0.33-0.72. The oligonucleotides are used to modulate
 CC expression of genes, particularly the genes for p53, ErbB-2, JunB, JunD,
 CC TGF-beta 1 or beta 2 to control proliferation of primary cell
 CC cultures (e.g. bone marrow stem, liver or kidney cells, osteoclasts,
 CC osteoblasts and/or keratinocytes). The oligonucleotides can also be
 CC used to analyse function of proteins (by altering their expression or
 CC activity) and therapeutically, e.g. in cases of cancer or (targeting
 CC TGF) for stimulating the immune system.
 SQ Sequence 14 BP; 2 A; 4 C; 3 G; 5 T;

alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-1 x V48937 ..

Align seg 1/1 to: V48937 from: 1 to: 14

1 Arg*****Leu 4
 |||||
 1 CGTAGTACTCTT 12

seq_name: N_Geneseq_36:V48608

seq_documentation_block:

ID V48608 standard; DNA; 14 BP.

AC V48608;

DT 15-OCT-1998 (first entry)

DE JunD gene antisense oligonucleotide JunD-25.

KW JunB; JunD; antisense oligonucleotide; modulate; gene expression; ss.

OS Synthetic.

OS Homo sapiens.

PN EP-856579-Al.

PD 05-AUG-1998.

PF 31-JAN-1997; 101531.

PR 31-JAN-1997; EP-101531.

PA (BIOG-) BIOGNOSTIK GES BIOMOLEKULARE DIAGNOSTIK.

PI Brysch W, Schlingensiefen K;

DR WPI: 98-400910/35.

PT Preparation of antisense oligo:nucleotide(s) which lack long runs of

PT consecutive guanosine or inosine - and have specific ratio of

PT residues able to form two or three hydrogen bonds, have greater

PT activity and reduced toxicity, used therapeutically or to modulate

PT growth of cells in culture

PS Claim 10: Fig 5a: 286pp; English.

CC V48564-708 represent antisense oligonucleotides directed against the

CC JunB and JunD genes. Of these, only oligonucleotides V48565-614 resulted

CC in effective downregulation of negative growth control by JunB or

CC JunD, while V48615-708 had little effect. The oligonucleotides

CC exemplify the invention. The specification describes oligonucleotides

CC that contain 8-30 nucleotides, which contain at most 8 nucleotides that

CC can each form three hydrogen bonds to cytosine; do not contain four

CC consecutive nucleotides able to form three H-bonds each to four

CC consecutive cytosines; do not contain two sequences of three consecutive

CC nucleotides each able to form three H-bonds to three consecutive

CC cytosines, and the ratio between residues able to form two H-bonds each

CC (2R) or three such bonds (3R) is given by 2R/3R = 0.33-0.72. The

CC oligonucleotides are used to modulate expression of genes, particularly

CC the genes for p53, ErbB-2, JunB, JunD, TGF-beta 1 or beta 2 to control

CC proliferation of primary cell cultures (e.g. bone marrow stem, liver or

CC kidney cells, osteoclasts, osteoblasts and/or keratinocytes). The

CC oligonucleotides can also be used to analyse function of proteins (by

CC altering their expression or activity) and therapeutically, e.g. in cases

CC of cancer or (targeting TGF) for stimulating the immune system.

SQ Sequence 14 BP; 1 A; 6 C; 4 G; 3 T;

alignment_scores:

Quality: 9.00 Length: 4

Ratio: 2.250 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-1 x V48608 ..

Align seg 1/1 to: V48608 from: 1 to: 14

1 Arg*****Leu 4
 |||||
 1 CGCTCCAGCTTG 12

seq_name: N_Geneseq_36:V48594

seq_documentation_block:

ID V48594 standard; DNA; 14 BP.

AC V48594;

DT 15-OCT-1998 (first entry)

DE JunD gene antisense oligonucleotide JunD-11.

KW JunB; JunD; antisense oligonucleotide; modulate; gene expression; ss.

OS Synthetic.

OS Homo sapiens.

PN EP-856579-Al.

PD 05-AUG-1998.

PF 31-JAN-1997; 101531.

PR 31-JAN-1997; EP-101531.

PA (BIOG-) BIOGNOSTIK GES BIOMOLEKULARE DIAGNOSTIK.

PI Brysch W, Schlingensiefen K;

DR WPI: 98-400910/35.

PT Preparation of antisense oligo:nucleotide(s) which lack long runs of

PT residues able to form two or three hydrogen bonds, have greater

PT activity and reduced toxicity, used therapeutically or to modulate

PT growth of cells in culture

PS Claim 10: Fig 5a: 286pp; English.

CC V48564-708 represent antisense oligonucleotides directed against the

CC JunB and JunD genes. Of these, only oligonucleotides V48565-614 resulted

CC in effective downregulation of negative growth control by JunB or

CC JunD, while V48615-708 had little effect. The oligonucleotides

CC exemplify the invention. The specification describes oligonucleotides

CC that contain 8-30 nucleotides, which contain at most 8 nucleotides that

CC can each form three hydrogen bonds to cytosine; do not contain four

CC consecutive nucleotides able to form three H-bonds each to four

CC consecutive cytosines; do not contain two sequences of three consecutive

CC nucleotides each able to form three H-bonds to three consecutive

CC cytosines, and the ratio between residues able to form two H-bonds each

CC (2R) or three such bonds (3R) is given by 2R/3R = 0.33-0.72. The

CC oligonucleotides are used to modulate expression of genes, particularly

CC the genes for p53, ErbB-2, JunB, JunD, TGF-beta 1 or beta 2 to control

CC proliferation of primary cell cultures (e.g. bone marrow stem, liver or

CC kidney cells, osteoclasts, osteoblasts and/or keratinocytes). The

CC oligonucleotides can also be used to analyse function of proteins (by

CC altering their expression or activity) and therapeutically, e.g. in cases

CC of cancer or (targeting TGF) for stimulating the immune system.

SQ Sequence 14 BP; 2 A; 4 C; 6 G; 2 T;

alignment_scores:

Quality: 9.00 Length: 4

Ratio: 2.250 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-1 x V48594/rev ..

Align seg 1/1 to reverse of: V48594 from: 1 to: 14

1 Arg*****Leu 4

Wed Feb 9 08:46:19 2000

us-08-653-294-1.rng

Page 7

|||||
14 CGCTCAGCCTG 3

THIS PAGE BLANK (USPTO)

FEATURES

```

source      1. .25
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human adult (K.Okubo)"
/dev_stage="adult"
BASE COUNT      5 a 7 c 5 g 8 t
ORIGIN

alignment_scores:
  Quality:      9.00      Length:      4
  Ratio:        2.250     Gaps:      0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x C21489/rev ..
Align seg 1/1 to reverse of: C21489 from: 1 to: 25

1 Arg*****Leu 4
|||||
19 AGGTCGACTCTA 8

seq_name: gb_est30:A1625681

seq_documentation_block:
LOCUS      A1625681      28 bp      mRNA      22-APR-1999
DEFINITION ty59c06.x1 NCI-CGAP Ut2 Homo sapiens cDNA clone IMAGE:2283370 3'
            similar to SW:COX1_HUMAN P00395 CYTOCHROME C OXIDASE POLYPEPTIDE I
            ; mRNA sequence.
ACCESSION  A1625681
VERSION    A1625681.1 GI:4650612
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 28)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    On Jun 5, 1998 this sequence version replaced gi:3187931.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert.Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Seq primer: -400P from Gibco
High quality sequence stop: 1.
FEATURES
  source      1. .28
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI-CGAP Ut2"
/tissue_type="moderately-differentiated endometrial
adenocarcinoma, 3 pooled tumors"
/lab_host="Dr10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.85 kb. Life Technologies catalog #:
11539-012"
BASE COUNT      7 a 5 c 9 g 7 t
ORIGIN

source      1. .28
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="Drosophila melanogaster P lethal line"
/note="Inverse PCR was performed on Drosophila
melanogaster strains each of which contains a single P
transposable element insertion that is thought to cause
either lethality or sterility. The resultant fragment for
each strain was directly sequenced to determine the
genomic sequence at the site of insertion. Details of the
protocols used can be found at
http://fruitfly.berkeley.edu/p-disrupt/inverse_pcr.html."
BASE COUNT      3 a 5 c 7 g 13 t
ORIGIN

alignment_scores:
  Quality:      9.00      Length:      4
  Ratio:        2.250     Gaps:      0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x A1625681/rev ..
Align seg 1/1 to reverse of: A1625681 from: 1 to: 28

1 Arg*****Leu 4
|||||
18 CGTACTACGTTG 7

seq_name: gb_gss8:AQ026240

seq_documentation_block:
LOCUS      AQ026240      28 bp      DNA      GSS      30-JUN-1998
DEFINITION l(3)L1231 Drosophila melanogaster P lethal line Drosophila
            melanogaster genomic sequence recovered from 5' end of P element,
            genomic survey sequence.
ACCESSION  AQ026240
VERSION    AQ026240.1 GI:3266525
KEYWORDS   GSS.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 28)
AUTHORS   Spradling,A.C., Stern,D., Beaton,A., Rehm,E.J., Lavery,T.,
            Mozdzen,N., Misra,S. and Rubin,G.M.
TITLE     The BDGP gene disruption project: Single P element insertions
            mutating 30% of Drosophila autosomal genes
JOURNAL    Unpublished (1998)
COMMENT    Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            FAX: 5106439947
            Email: gerry@fruitfly.berkeley.edu

Sequence recovery method was inverse PCR.

Sequence orientation is forward strand relative to 5' end of P
element

The P element insertion position is base 021 in the 28 bases. This
insertion position refers to the first base of the 8 base target
recognition sequence.

FEATURES
  Class: transposon-tagged
  Location/Qualifiers
    1. .28
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="Drosophila melanogaster P lethal line"
/note="Inverse PCR was performed on Drosophila
melanogaster strains each of which contains a single P
transposable element insertion that is thought to cause
either lethality or sterility. The resultant fragment for
each strain was directly sequenced to determine the
genomic sequence at the site of insertion. Details of the
protocols used can be found at
http://fruitfly.berkeley.edu/p-disrupt/inverse_pcr.html."
BASE COUNT      3 a 5 c 7 g 13 t
ORIGIN

```

Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-1 x AQ026240/rev ..

Align seg 1/1 to reverse of: AQ026240 from: 1 to: 28

1 Arg*****Leu 4

|||||

15 CGAAGTACGTTG 4

seq_name: gb_gss13:AQ051551

seq_documentation_block:

LOCUS AQ051551 28 bp DNA GSS 24-MAR-1999
DEFINITION nbxb0002bh03r CUGI Rice BAC Library Oryza sativa genomic clone
nbxb0002006r, genomic survey sequence.

ACCESSION AQ051551

VERSION AQ051551.2 GI:4501321

KEYWORDS GSS.

SOURCE Oryza sativa.

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Oryza.

REFERENCE 1 (bases 1 to 28)

AUTHORS Wing, R.A. and Dean, R.A.

TITLE A BAC End Sequencing Framework to Sequence the Rice Genome

JOURNAL Unpublished (1998)

COMMENT On Mar 23, 1999 this sequence version replaced gi:3324109.

Contact: Wing RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Seq primer: GGAAACGCTATGACCATG

Class: BAC ends

High quality sequence stop: 1.

FEATURES

Source

1..28

/organism="Oryza sativa"

/strain="Japonica"

/cultivar="Nipponbare"

/db_xref="taxon:4530"

/clone="nbxb0002006r"

/clone_lib="CUGI Rice BAC Library"

/tissue_type="leaf"

/lab_host="E. coli DH10B"

/note="Vector: pBel0BAC11; Site_1: HindIII; Site_2:
HindIII. Rice is one of two most popular grains in the
world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 Mbp
(Arumuganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from Oryza sativa, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 Kb providing 10.9 haploid genome equivalents.
The deep coverage allows the isolation a particular
sequence with a probability of 99.9 %. Two high density
filters, each containing 18,432 clones (doubly spotted),
represent the whole library for colony screening."

BASE COUNT

ORIGIN

9 a 7 c 8 g 4 t

alignment_scores:

Quality: 9.00

Ratio: 2.250

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-1 x AQ051551/rev ..

Align seg 1/1 to reverse of: AQ051551 from: 1 to: 28

1 Arg*****Leu 4

|||||

20 CGTACTACACTC 9

seq_name: gb_est6:N85094

seq_documentation_block:

LOCUS N85094 29 bp mRNA EST 01-APR-1996
DEFINITION J2291F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA
clone J2291 5' similar to RAB7P, mRNA sequence.

ACCESSION N85094

VERSION N85094.1 GI:1260719

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 29)

AUTHORS Liew, C.C.

TITLE cDNAs from fetal heart (1996)

JOURNAL Unpublished (1996)

COMMENT Contact: Liew CC

Department of Laboratory Medicine and Pathobiology

University of Toronto

Banting Institute, 100 College St., Toronto, Ontario, M5G1L5

Tel: 416/9788758

Fax: 416/9788550

Email: liewcc@utcc.utoronto.ca

Seq primer: GAAATTAACCTCACTAAAGGG.

FEATURES

Source

1..29

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="J2291"

/clone_lib="Human fetal heart, Lambda ZAP Express"

/lab_host="E. coli XL1-Blue"

/note="Vector: Lambda ZAP Express; Site_1: EcoRI; Site_2:
XhoI; mRNA was purified from human fetal hearts (8-10
weeks). cDNA was synthesized using a XhoI-Oligo dT
adaptor-primer. EcoRI adaptors were ligated, followed by
digestion with XhoI, for directional cloning into
predigested lambda ZAP Express."

BASE COUNT

ORIGIN

8 a 6 c 10 g 5 t

alignment_scores:

Quality: 9.00

Ratio: 2.250

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-1 x N85094 ..

Align seg 1/1 to: N85094 from: 1 to: 29

1 Arg*****Leu 4

|||||

15 AGAACTAGTCTC 26

seq_name: gb_est5:D18723


```

seq_documentation_block:
LOCUS D18723 30 bp mRNA EST 12-DEC-1995
DEFINITION MUSGS01785 Mouse 3'-directed Mus musculus domesticus cDNA clone
md0188 3', mRNA sequence.
ACCESSION D18723
VERSION D18723.1 GI:1100692
KEYWORDS EST.
SOURCE western European house mouse.
ORGANISM Mus musculus domesticus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 30)
AUTHORS Kawamoto,S., Okubo,K., Yoshii,J., Katsuki,M. and Matsubara,K.
TITLE Analysis of gene expression in mouse embryogenesis by 3'-directed
cDNA sequencing
JOURNAL Unpublished (1995)
COMMENT On May 18, 1995 this sequence version replaced gi:811444.
Contact: Kawamoto,S., Okubo,K., Yoshii,J., Katsuki,M. and
Matsubara,K.
Institute for Cellular and Molecular Biology
Osaka University
3-1 Yamada-Oka, Suita, Osaka 565, Japan
Insert Length: 783 Std Error: 0.00
High quality sequence stop: 354.
FEATURES
Location/Qualifiers
BASE COUNT 9 a 12 c 4 g 5 t
ORIGIN
1..30
/organism="Mus musculus domesticus"
/strain="CS7BL/6J"
/db_xref="taxon:10092"
/clone="md0188"
/clone_lib="Mouse 3'-directed"
/tissue_type="decidual tissue (day 6.5-8.5 of gestation)"

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x D18723/rev ..
Align seg 1/1 to reverse of: D18723 from: 1 to: 30

1 Arg*****Leu 4
|||||
19 AGGTCGACTCTA 8

seq_name: gb_est8:C01074

seq_documentation_block:
LOCUS C01074 30 bp mRNA EST 23-JUL-1996
DEFINITION HUMGS0007722 Human adult (K.Okubo) Homo sapiens cDNA, mRNA
sequence.
ACCESSION C01074
VERSION C01074.1 GI:1433304
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 30)
AUTHORS Okubo,K.
TITLE BodyMap: human gene expression database
JOURNAL Unpublished (1995)
COMMENT Contact: Okubo,K.
Institute for Molecular and Cellular Biol
Osaka University
1-3,Yamada-Oka, Suita, Osaka Pref. 565, Japan
Tel: 06-877-5111(ex.3315)
Email: kousaku@imcb.osaka-u.ac.jp

```

```

Human Gene Signature, 3'-directed cDNA sequence. We are not
submitting the same cDNA sequence redundantly to DDBJ since 1993.
For the abundance information of clones with this sequence in this
library and as well as in other 3'-directed libraries, see
http://www.imcb.osaka-u.ac.jp/bodymap/. The sequences of the clones
represented by this GS sequences is also found there.
FEATURES
Location/Qualifiers
BASE COUNT 9 a 6 c 6 g 9 t
ORIGIN
1..30
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human adult (K.Okubo)"
/dev_stage="adult"

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x C01074/rev ..
Align seg 1/1 to reverse of: C01074 from: 1 to: 30

1 Arg*****Leu 4
|||||
28 AGGGCATCACTT 17

seq_name: gb_gss8:AQ025148

seq_documentation_block:
LOCUS AQ025148 30 bp DNA GSS 14-OCT-1998
DEFINITION EP(3)1001 Drosophila melanogaster EP line Drosophila melanogaster
genomic Sequence recovered from 5' end of P element, genomic survey
sequence.
ACCESSION AQ025148
VERSION AQ025148.1 GI:3265500
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 30)
AUTHORS Rehm,E.J. and Rubin,G.M.
TITLE The BDGP gene disruption project: single EP element insertions
JOURNAL Unpublished (1998)
COMMENT Contact: Gerald Rubin
Berkeley Drosophila Genome Project
University of California, Berkeley
LSA Building, Berkeley, CA 94720-3200, USA
Fax: 5106433947
Email: gerry@fruitfly.berkeley.edu

Sequence recovery method was inverse PCR.
Sequence orientation is forward strand relative to 5' end of P
element

The P element insertion position is base 23 in the 30 bases. This
insertion position refers to the first base of the 8 base target
recognition sequence.
Class: transposon-tagged.
Location/Qualifiers
BASE COUNT 1..30
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="Drosophila melanogaster EP line"
/note="Inverse PCR was performed on Drosophila
melanogaster strains each of which contains a single EP

```

transposable element insertion. (The generation of these insertion strains is described in Rorth P, Szabo K, Bailey A, Lavery T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function genetics in Drosophila. Development 6:1049-1057.) The resultant fragment for each strain was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html.

BASE COUNT
ORIGIN

0 a 11 c 10 g 9 t

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x A0025148 ..

Align seg 1/1 to: A0025148 from: 1 to: 30

1 Arg*****Leu 4
|||||||

2 CGCTCGCTCTC 13

seq_name: gb_est1:T48676

seq_documentation_block:

LOCUS T48676 31 bp mRNA EST 06-FEB-1995
DEFINITION Yb01f01.s1 Stratagene placenta (#937225) Homo sapiens cDNA clone IMAGE:69913 3' similar to gb:J00118 LACTOGEN PRECURSOR (HUMAN), mRNA sequence.

ACCESSION T48676
VERSION T48676.1 GI:650536
KEYWORDS human.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 31)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chapelli,B., Chisoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.

TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Other ESTs: yb01f01.r1

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

Insert Size: 362
High quality sequence starts: 1 High quality sequence stops: 1
Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Trace considered overall poor quality
Insert Length: 362 Std Error: 0.00
Seq primer: -21m13

High quality sequence stop: 1.
Location/Qualifiers

FEATURES
source

1..31
/organism="Homo sapiens"
/db_xref="GDB:490810"
/db_xref="taxon:9606"
/clone="IMAGE:69913"

/clone_lib="Stratagene placenta (#937225)"
/sex="male"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: Vector: pBluescript SK-; Site: 1.
EcoRI; Site: 2. XhoI; Cloned unidirectionally. Primer:
Oligo dt. Caucasian. Average insert size: 1.2 kb; Uni-TAP
XR Vector; -5' adaptor sequence: 5' GAATTCGGCAG 3' -3'
adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3'"
BASE COUNT 8 a 13 c 6 g 3 t 1 others
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x T48676 ..

Align seg 1/1 to: T48676 from: 1 to: 31

1 Arg*****Leu 4
|||||||

15 AGACGGGCACTG 26

seq_name: gb_est5:D18720

seq_documentation_block:

LOCUS D18720 31 bp mRNA EST 12-DEC-1995
DEFINITION MUSGS01782 Mouse 3'-directed Mus musculus domesticus CDNA clone md0435 3', mRNA sequence.

ACCESSION D18720
VERSION D18720.1 GI:1100689
KEYWORDS EST.
SOURCE western European house mouse.

ORGANISM Mus musculus domesticus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 31)
AUTHORS Kawamoto,S., Okubo,K., Yoshi,J., Katsuki,M. and Matsubara,K.
TITLE Analysis of gene expression in mouse embryogenesis by 3'-directed cDNA sequencing
JOURNAL Unpublished (1995)
COMMENT On May 18, 1995 this sequence version replaced gi:811441.
Contact: Kawamoto,S., Okubo,K., Yoshi,J., Katsuki,M. and Matsubara,K.

Institute for Cellular and Molecular Biology
Osaka University
3-1 Yamada-oka, Suita, Osaka 565, Japan
Insert Length: 694 Std Error: 0.00
High quality sequence stop: 370.

FEATURES
Location/Qualifiers

1..31
/organism="Mus musculus domesticus"
/strain="C57BL/6J"
/db_xref="taxon:10092"
/clone="md0435"
/clone_lib="Mouse 3'-directed"
/tissue_type="decidual tissue (day 6.5-8.5 of gestation)"

BASE COUNT 9 a 7 c 5 g 10 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-1 x D18720/rev ..

Align seg 1/1 to reverse of: D18720 from: 1 to: 31

1 Arg*****Leu 4
|||||
19 AGGTCGACTCTA 8

seq_name: gb_est22:AI021071

seq_documentation_block: 31 bp mRNA 16-JUN-1998
LOCUS AI021071
DEFINITION ua99f03.r1 Soares mouse mammary gland NBMWG Mus musculus cDNA clone IMAGE:1365633 5' similar to SW:PERL_RAT P43884 PERILIPIN A/B ; mRNA sequence.

ACCESSION AI021071
VERSION AI021071.1 GI:3235407
KEYWORDS EST.
SOURCE house mouse.

ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 31)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2150979.

CONTACT: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:898853

Trace considered overall poor quality
Possible reversed clone; similarity on wrong strand
Seq primer: -28ml3 rev2 Et from Amersham
High quality sequence stop: 1.

FEATURES

source
1..31
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1365633"
/clone_lib="Soares mouse mammary gland NBMWG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCGAATGTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 3 a 8 c 16 g 4 t
ORIGIN

alignment_scores:

Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-1 x AI021071/rev

Align seg 1/1 to reverse of: AI021071 from: 1 to: 31

1 Arg*****Leu 4
|||||
21 CGCACCTCACTC 10

seq_name: gb_est22:AI032352

seq_documentation_block: 31 bp mRNA 23-JUN-1998
LOCUS AI032352
DEFINITION ou48d08.x1 NCI_CGAP_Br2 Homo sapiens cDNA clone IMAGE:1631055 3' similar to TR:Q62098 Q62098 PUTATIVE PRIMORDIAL PROTEIN TRANSCRIPT ; mRNA sequence.

ACCESSION AI032352
VERSION AI032352.1 GI:3250564
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 31)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2151514.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Seq primer: -40ml3 fwd. Et from Amersham

High quality sequence stop: 1.

FEATURES

source
1..31
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1631055"
/clone_lib="NCI_CGAP_Br2"
/sex="female, pooled"
/tissue_type="breast"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled bulk breast tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. This library is the normalized version of NCI_CGAP_Br1.1. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 8 a 10 c 1 g 12 t
ORIGIN

alignment_scores:

Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-1 x AI032352

Align seg 1/1 to: AI032352 from: 1 to: 31

```
1 Arg*****Leu 4
  |||||
  6 CGATCTACACTA 17
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:29:33 : Search time 122.56 Seconds
(without alignments)
1.933 Million cell updates/sec

Title: US-08-653-294-2
Perfect score: 17
Sequence: 1 XXXXXLXXR 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|----------|--------------------|
| 1 | 9 | 52.9 | 4 | 1 R14267 | Chromogenic substr |
| 2 | 9 | 52.9 | 5 | 1 P40822 | Sequence of fragme |
| 3 | 9 | 52.9 | 5 | 1 W00248 | Subtilisin N62D/G1 |
| 4 | 9 | 52.9 | 6 | 1 R57725 | Rheumatoid arthrit |
| 5 | 9 | 52.9 | 6 | 1 W48376 | Protein tyrosine k |
| 6 | 9 | 52.9 | 6 | 1 W93209 | Activin receptor p |
| 7 | 9 | 52.9 | 7 | 1 R37709 | Deltal4 ser17 hCNT |
| 8 | 9 | 52.9 | 7 | 1 R71101 | PTK catalytic doma |
| 9 | 9 | 52.9 | 7 | 1 W11186 | Brain homing pepti |
| 10 | 9 | 52.9 | 7 | 1 W49319 | CLK serine/threoni |
| 11 | 9 | 52.9 | 7 | 1 W50983 | Growth hormone rel |
| 12 | 9 | 52.9 | 7 | 1 W81413 | PTK kinase domain |
| 13 | 9 | 52.9 | 7 | 1 Y04013 | Peptide inhibiting |
| 14 | 9 | 52.9 | 8 | 1 R36108 | Hepatitis C virus |
| 15 | 9 | 52.9 | 8 | 1 R36109 | Hepatitis C virus |
| 16 | 9 | 52.9 | 8 | 1 R36110 | Hepatitis C virus |
| 17 | 9 | 52.9 | 8 | 1 R69855 | Rheumatoid arthrit |
| 18 | 9 | 52.9 | 8 | 1 R63089 | PTK subdomain IV. |
| 19 | 9 | 52.9 | 8 | 1 W75839 | Mouse mast cell pr |
| 20 | 9 | 52.9 | 8 | 1 W75808 | Mouse mast cell pr |
| 21 | 9 | 52.9 | 8 | 1 W64270 | mMCP-7 peptide sub |
| 22 | 9 | 52.9 | 8 | 1 Y00373 | Fragment of human |
| 23 | 9 | 52.9 | 9 | 1 R59247 | Peptide fragment (|
| 24 | 9 | 52.9 | 9 | 1 R80215 | Protein tyrosine k |
| 25 | 9 | 52.9 | 9 | 1 R89172 | Peptide HER2/neu C |
| 26 | 9 | 52.9 | 9 | 1 R98352 | PK1 subdomain TK6 |
| 27 | 9 | 52.9 | 9 | 1 R99915 | Hepatitis C Virus |
| 28 | 9 | 52.9 | 9 | 1 W13412 | Brain homing pepti |
| 29 | 9 | 52.9 | 9 | 1 W49303 | Human leucocyte an |
| 30 | 9 | 52.9 | 9 | 1 W79626 | Truncated GHRH ana |
| 31 | 9 | 52.9 | 9 | 1 W79610 | Truncated GHRH ana |
| 32 | 9 | 52.9 | 10 | 1 P40094 | Sequence at antige |
| 33 | 9 | 52.9 | 10 | 1 W32746 | Human platelet gly |
| 34 | 9 | 52.9 | 10 | 1 W13715 | Internal neururin |

| | | | | | |
|----|---|------|----|----------|--------------------|
| 35 | 9 | 52.9 | 10 | 1 W30080 | Hamster neururin |
| 36 | 9 | 52.9 | 10 | 1 W79609 | Truncated GHRH ana |
| 37 | 9 | 52.9 | 10 | 1 W79634 | Truncated GHRH ana |
| 38 | 9 | 52.9 | 10 | 1 W79619 | Truncated GHRH ana |
| 39 | 9 | 52.9 | 11 | 1 W23143 | NGF receptor p75-i |
| 40 | 9 | 52.9 | 11 | 1 W79608 | Truncated GHRH ana |
| 41 | 9 | 52.9 | 11 | 1 W79628 | Truncated GHRH ana |
| 42 | 9 | 52.9 | 11 | 1 W79623 | Truncated GHRH ana |
| 43 | 9 | 52.9 | 11 | 1 Y04022 | Cyclic peptide inh |
| 44 | 9 | 52.9 | 12 | 1 P61068 | Sequence of antige |
| 45 | 9 | 52.9 | 12 | 1 R37711 | Glyp14Ser17 CNTF N |

ALIGNMENTS

| | |
|-------------------------|--|
| RESULT 1 | |
| R14267 | |
| ID | R14267 standard; peptide; 4 AA. |
| AC | R14267; |
| DT | 02-JAN-1992 (first entry) |
| DE | Chromogenic substrate peptide. |
| KW | Identification; specific enzymatic cleavage. |
| OS | Synthetic. |
| FT | Key Location/Qualifiers |
| FT | modified_site 4 |
| FT | /tag= a |
| FT | /note= "Arg-pNa, pNa = p-nitroanilide" |
| PN | W09114787-A. |
| PD | 03-OCT-1991. |
| PF | 22-MAR-1991; F00233. |
| PR | 23-MAR-1990; FR-003726. |
| PA | (SERB-) SERBIO. |
| PI | Lepargneur JP, Pussard-Contant G, Martinoli JL, Quentin G; |
| DI | WPI: 91-310586/42 |
| PT | Identification of Candida and Torulopsis species - by |
| PT | Species-specific enzymatic cleavage of chromogenic substrates |
| PS | Claim 5; Page 43; 48pp; French. |
| CC | The chromogenic substrate peptide is used in the identification |
| CC | of the nonpathogenic Candida species C.guilliermondii. The peptide |
| CC | has the formula Boc-Leu-Ser-Thr-Arg-pNa. Identification is |
| CC | effected by incubating the sample with the peptide which is |
| CC | cleavable with at least one aminoamidase or aminopeptidase and |
| CC | observing the kinetics of the cleavage induced colour formation. |
| CC | This process may be used to distinguish pathogenic Candida and |
| CC | Torulopsis species from nonpathogenic species. |
| SQ | Sequence 4 AA; |
| Query Match | 52.9%; Score 9; DB 1; Length 4; |
| Best Local Similarity | 50.0%; Pred. No. 1.5e-05; |
| Matches 2; Conservative | 0; Mismatches 2; Indels 0; Gaps 0; |
| QY | 7 LXXR 10 |
| Db | 1 LSTR 4 |
| RESULT 2 | |
| P40822 | |
| ID | P40822 standard; peptide; 5 AA. |
| AC | P40822; |
| DT | 03-AUG-1992 (first entry) |
| DE | Sequence of fragment F corresp. to residues 16-20 of human |
| DE | pancreatic growth hormone releasing factor (somatocorin) (hpGRF). |
| KW | Hormone; dwarfism; therapy; retarded growth; |
| KW | anabolic protein deficiency; growth promoter; lactation. |
| OS | Homo sapiens. |
| FT | Key Location/Qualifiers |
| FT | modified_site 1 |
| FT | /label= Boc-O |
| FT | /note= "Boc-tertobutyloxycarbonyl (carbamate)" |
| FT | modified_site 5 |

```

FT AU8424774-A.
PN 30-AUG-1984.
PD 20-FEB-1984.
PF 400343.
PR 21-FEB-1983; FR-002781.
PR 29-NOV-1983; FR-019058.
PA (SNFI) SANOFI SA.
PI Diaz J, Demarne H, Roncucci R, Schmelck PH;
DR WPI; 84-256760/42.
PT Synthesis of hpGRF in liquid phase reactions - with use of new
PT peptide fragments
PS Claim 4; page 67; 76pp; French.
CC The inventors claim hpGRF fragments used for the synthesis of hpGRF.
CC Using the method somatostatin can be obtd. on a large scale with
CC good yield and good purity. The prod. is used in man for the
CC treatment of dwarfism and retarded growth and for anabolic protein
CC deficiencies. In animals it is useful for promoting wt. growth for
CC increased prodn., lactation etc. This index is based on EP-122818,
CC which is the equivalent of AU8424774.
SQ Sequence 5 AA;

Query Match 52.9%; Score 9; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 1.5e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
DB 2 LSAR 5

RESULT 3
W00248
ID W00248 standard; Peptide: 5 AA.
AC W00248;
DT 16-DEC-1996 (first entry)
DE Subtilisin N62D/G166D substrate peptide.
KW Subtilisin BPN'; protease; enzyme engineering; protein engineering;
KW Bacillus amyloliquefaciens; mutagenesis; substrate specificity;
KW phage display.
OS Synthetic.
PN W09627671-A1.
PD 12-SEP-1996.
PF 27-FEB-1996; U02861.
PR 03-MAR-1995; US-398028.
PR 19-JUL-1995; US-504265.
PA (GETH) GENENTECH INC.
PI Ballinger MD, Wells JA;
DR WPI; 96-425431/42.
PT Subtilisin variants for cleaving substrates contg. basic residues -
PT allow effective cleavage of fusion proteins with basic linker
sequences
PS Example 4; Page 27; 83pp; English.
CC A phage display system was used to screen 5-residue substrate
CC linkers for cleavage by subtilisin BPN' N62D/G166D double mutant
CC (W00247). The library was subjected to 9 rounds of selection, and
CC clones that were increasingly sensitive or resistant to cleavage
CC were selected. Of 21 clones in the sensitive pool, 3 contained a
CC monobasic substrate linker sequence (W00248) contg. a hydrophobic
CC residue at P4; the remaining 18 were dibasic (see also W00249-53).
CC Of 10 substrates selected from the protease resistant pool (W00254-
CC 63), 7 contained no basic sites, 2 were monobasic and 1 was dibasic.
CC The N62D/G166D double mutant specifically cleaves protein substrates
CC contg. basic amino acid residues at positions P1 and P2.
SQ Sequence 5 AA;

Query Match 52.9%; Score 9; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 1.5e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
DB 2 LSAR 5

/label= R-OH

FT AU8424774-A.
PN 30-AUG-1984.
PD 20-FEB-1984.
PF 400343.
PR 21-FEB-1983; FR-002781.
PR 29-NOV-1983; FR-019058.
PA (SNFI) SANOFI SA.
PI Diaz J, Demarne H, Roncucci R, Schmelck PH;
DR WPI; 84-256760/42.
PT Synthesis of hpGRF in liquid phase reactions - with use of new
PT peptide fragments
PS Claim 4; page 67; 76pp; French.
CC The inventors claim hpGRF fragments used for the synthesis of hpGRF.
CC Using the method somatostatin can be obtd. on a large scale with
CC good yield and good purity. The prod. is used in man for the
CC treatment of dwarfism and retarded growth and for anabolic protein
CC deficiencies. In animals it is useful for promoting wt. growth for
CC increased prodn., lactation etc. This index is based on EP-122818,
CC which is the equivalent of AU8424774.
SQ Sequence 5 AA;

Query Match 52.9%; Score 9; DB 1; Length 6;
Best Local Similarity 50.0%; Pred. No. 1.5e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
DB 1 LTR 4

RESULT 4
R57725
ID R57725 standard; peptide; 6 AA.
AC R57725;
DT 18-MAR-1995 (first entry)
DE Rheumatoid arthritis-predictive factor.
KW Rheumatoid arthritis; diagnosis; therapy; fibronectin; streptokinase.
OS Synthetic.
PN W09417411-A.
PD 04-AUG-1994.
PF 27-JAN-1994; U01077.
PR 27-JAN-1993; US-009471.
PA (TRIN-) TRINITY LAB INC.
PA (UYDU-) UNIV DUKE.
PI Clinton BK, Gonzalez-Gronow M, Pizzo SV;
DR WPI; 94-264260/32;
PT Diagnosis and treatment of rheumatoid arthritis - using a
PT predictive factor peptide corresp. to a region common to
PT fibronectin and streptokinase
PS Claim 1; Page 54; 83pp; English.
CC The peptide corresponds to a common region/epitope shared by
CC streptokinase (SK) and fibronectin that is recognized by anti-SK
CC antibodies taken from rheumatoid arthritis (RA) patients. The
CC peptide is used in the diagnosis of the onset/presence of RA, for
CC monitoring the progress of therapy, or for therapy.
SQ Sequence 6 AA;

Query Match 52.9%; Score 9; DB 1; Length 6;
Best Local Similarity 50.0%; Pred. No. 1.5e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
DB 1 LTR 4

RESULT 5
W48376
ID W48376 standard; Peptide; 6 AA.
AC W48376;
DT 17-AUG-1998 (first entry)
DE Protein tyrosine kinase conserved peptide.
KW RPTK; related adhesion focal tyrosine kinase; human;
KW protein tyrosine kinase; PTK; cell growth; cell differentiation;
KW cell adhesion; cell migration.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Misc_difference 6
FT /label= Phe, Tyr
PN W09807870-A1.
PD 26-FEB-1998.
PF 12-AUG-1997; U14093.
PR 13-MAR-1997; US-816462.
PR 23-AUG-1996; US-703623.
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PI Abraham H, Avraham S, Groopman JB;
DR WPI; 98-169174/15.
PT New nucleic acid encoding vertebrate related adhesion focal tyrosine
PT kinase - useful for, e.g. regulating growth, differentiation,
PT adhesion and migration of cells, used for treating metastases
PS Example 1; Page 72; 167pp; English.
CC This peptide represents a conserved region of protein tyrosine
CC kinases (PTKs). A primer (see V17802) based on the peptide was
CC used with another primer (see V17803) also based on a conserved
CC peptide to amplify cDNA from human megakaryocytic leukaemia CMK
CC cells. A novel 160 bp product was obtained, and was in turn used
CC to screen a human hippocampal cDNA library for identification of
CC additional clones. The complete coding region (see V17800) for a
CC novel PTK (see W48373), designated vertebrate related adhesion

```

CC focal tyrosine kinase (RAFTK), was assembled from 8 overlapping
 CC clones.
 SQ Sequence 6 AA;

Query Match 52.9%; Score 9; DB 1; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 |
 Db 2 LAAR 5

RESULT 6
 W93209
 ID W93209 standard; Protein; 6 AA.
 AC W93209;
 DT 27-MAY-1999 (first entry)
 DE Activin receptor predictive kinase domain #5.
 KW Activin receptor; activin/TGF-superfamily; receptor protein; function;
 KW activity; modulate; treatment; carcinogenesis; wound healing; fertility;
 KW immune system disorder; central nervous system disorder; kinase domain;
 KW reproductive system.
 OS Unidentified.
 PN US885794-A.
 PD 23-MAR-1999.
 PF 02-SEP-1994; 300584.
 PR 08-MAY-1992; US-880220.
 PR 10-MAY-1991; US-698709.
 PR 09-OCT-1991; US-773229.
 PR 02-SEP-1994; US-300584.
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 PI Mathews LS, Valle WW;
 DR WPI; 99-228534/19.
 PT Isolated nucleic acid molecules encoding vertebrate activin receptor
 PT polypeptides - useful as probes for detecting similar sequences and
 PT for investigating the function of the receptor in conditions such as
 PT carcinogenesis, wound healing and disorders of the immune, central
 PT nervous and reproductive systems
 PS Example VIII; Column 35; 28pp; English.
 CC This sequence represents an activin receptor polypeptide predictive
 CC kinase domain motif. The nucleic acid molecules of the invention are
 CC useful as probes for the identification of additional members of the
 CC activin/TGF-superfamily of receptor proteins, and the coding sequences
 CC can be used for the recombinant expression of the receptor proteins or
 CC functional fragments of them. They may also be used to study the function
 CC and activity of activin receptor polypeptides in cells and to identify
 CC agents which will modulate activin receptor expression and activity for
 CC use in treating conditions such as carcinogenesis, wound healing,
 CC disorders of the immune or central nervous systems and especially the
 CC reproductive system (where they may be used to control fertility in
 CC humans, domestic and commercial animals).
 SQ Sequence 6 AA;

Query Match 52.9%; Score 9; DB 1; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 |
 Db 2 LAAR 5

RESULT 7
 R37709
 ID R37709 standard; Peptide; 7 AA.
 AC R37709;
 DT 28-SEP-1993 (first entry)
 DE Delta4 Ser17 hCNTF primer design peptide #2.
 KW Human; ciliary neurotrophic factor; hCNTF; intrinsic muscle; eye;
 KW choroid; ciliary body; iris; growth; chromaffin adrenal cell; CHAT;

KW chick; retinal cell; cholinergic neuron; pontine basal region; brain;
 KW cholinergic neurotrophic factor; sympathetic neuron; spinal neuron;
 KW cranial ganglia; motoneuron; spinal cord; hippocampal neuron; 0-2A;
 KW differentiation; progenitor; glial cell; oligodendrocyte; astrocyte;
 KW ganglioside; nervous disorders; PCR; amplify; primer; pT7.7.
 OS Synthetic.
 PN WO9310233-A.
 PD 27-MAY-1993.
 PF 11-NOV-1992; E02386.
 PR 11-NOV-1991; IT-PD0207.
 PR 20-MAY-1992; IT-PD0087.
 PA (FIDI-) FIDIA SPA.
 PI Callegaro L, Negro A;
 DR WPI; 93-182549/22.
 DR N-PSDB; 042974.
 PT Truncated and/or mutein human ciliary neurotrophic factor -
 PT useful for treating neuro-pathological disorders due to loss of
 PT nervous function
 PS Disclosure; Page 27; 66pp; English.
 CC This sequence was used in the design of a primer which was used in the
 CC production of a human ciliary neurotrophic factor (hCNTF) analogue
 CC lacking the N-terminal 14 amino acids and Cys17 replaced by Ser. This
 CC modification reduces the formation of CNTF dimers thereby maintaining
 CC biological activity. This primer contains a BglII site in the 5' end
 CC and the seventeenth codon is changed to TCT. The amplified sequence
 CC was ligated into plasmid pSERTB and transcribed. Wild-type CNTF
 CC enhances the survival and development of some sympathetic neurons,
 CC spinal neurons of cranial ganglia, motoneurons of the spinal cord and
 CC hippocampal neurons. It has been shown to influence the different-
 CC iation of the progenitor glial cells known as 0-2A which differentiate
 CC in oligodendrocytes and astrocytes. Truncated forms of CNTF are useful
 CC in a pharmaceutical composition, which may also comprise a natural
 CC ganglioside or derivative or a semisynthetic analogue, for treating
 CC nervous disorders. The truncated proteins maintain, prevent loss of
 CC or recover nervous function. They can be used to treat neuropath-
 CC ological conditions caused by aging of the nervous system or diseases
 CC of the immune system.
 SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 |
 Db 2 LSSR 5

RESULT 8
 R71101
 ID R71101 standard; Peptide; 7 AA.
 AC R71101;
 DT 17-AUG-1995 (first entry)
 DE PTK catalytic domain conserved region.
 KW Protein-tyrosine-kinase; PTK; discoidin domain receptor; cancer;
 KW breast tumor; mamma carcinoma; diagnosis; prognosis; therapy.
 OS Homo sapiens.
 PN WO9502187-A.
 PD 19-JAN-1995.
 PF 08-JUL-1994; G01480.
 PR 09-JUL-1993; GB-014271.
 PA (CANC-) CANCER RES INST.
 PA (WELL) WELLCOME FOUND LTD.
 PI Barker KT, Crompton MR, Gusterson BA, Martindale JE;
 PI Mitchell PJ, Page MJ, Spence P;
 DR WPI; 95-066991/09.
 PT Method for screening substances, using protein tyrosine kinase -
 PT for potential utility as therapeutic agents for cancer
 PS Disclosure; Page 33; 51pp; English.
 CC CDNA derived from tumor metastatic tissue was amplified using
 CC primers (given in Q84783-84) based on sequences (R71101, R71103)
 CC associated with protein-tyrosine-kinases (PTK). Novel PTK22 was

CC identified in an isolated subclone. The 3' sequence of PTK22 was
 CC obtained by reverse transcription (using the primer of Q84786) and
 CC PCR amplification (primers Q84787-88) of RNA of human breast
 CC carcinoma cell line MDA MG 468. The partial DNA sequence of PTK22
 CC is given in Q84782.
 SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 LXXR 10
 | |
 Db 3 LAAR 6

RESULT 9

W11186
 ID W11186 standard; Peptide; 7 AA.
 AC W11186;
 DT 15-JAN-1998 (first entry)
 DE Brain homing peptide.
 KW Brain homing peptide; in vivo panning; screening; phage display;
 KW drug delivery.
 OS Synthetic.
 PN W09710507-A1.
 PD 20-MAR-1997.
 PF 10-SEP-1996; U14600.
 PR 11-SEP-1995; US-526710.
 PR 11-SEP-1995; US-526708.
 PA (LJOL-) LA JOLLA CANCER RES FOUND.
 PI Pasqualini R, Ruoslahti E;
 DR WPI; 97-202359/18.
 PT Obtaining compound that homes to selected organ or tissue - by in
 PT vivo panning method, specifically to identify brain, kidney,
 PT angiogenic vasculature or tumour tissue homing peptide(s)
 PS Disclosure; Page 45; 75pp; English.
 CC This synthetic peptide is an example of a brain-homing peptide
 CC that was identified using a claimed method for obtaining
 CC molecules that home to a selected organ or tissue. This in vivo
 CC panning method typically involves administering a phage display
 CC library to a subject, and identifying expressed peptides which
 CC home to the desired organ or tissue, e.g. brain, kidney, angiogenic
 CC vascular tissue or tumour tissue. The isolated peptides (see
 CC W13412-52, W11181-86) can be used to target e.g. drugs, toxins or
 CC labels to the selected organ/tissue (claimed) or to identify and/or
 CC isolate target molecules (claimed). The peptides can be directly
 CC identified in vivo, as compared to prior art in vitro screening
 CC methods, which require further examination to see if they maintain
 CC specificity in vivo.
 SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 LXXR 10
 | |
 Db 1 LSSR 4

RESULT 10

W49919
 ID W49919 standard; Peptide; 7 AA.
 AC W49919;
 DT 20-JUL-1998 (first entry)
 DE CLK serine/threonine kinase consensus peptide.
 KW CLK serine/threonine kinase; protein kinase; mCLK1; mCLK2; mCLK3;
 KW mCLK4; mouse; signal transduction.
 OS Synthetic.
 PN W09748723-A2.

PD 24-DEC-1997.
 PF 17-JUN-1997; IB0946.
 PR 19-DEC-1996; US-034286.
 PR 17-JUN-1996; US-019629.
 PR 09-AUG-1996; US-023485.
 PR 13-NOV-1996; US-030860.
 PR 15-NOV-1996; US-030964.
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PI Aoki N, Chen Z, Kharitonkov AI, Kim YW, Nayler O,
 PI Ullrich A, Wang HY;
 DR WPI; 98-120302/11.
 PT New phosphatase and kinase enzyme(s) - useful in the diagnosis and
 PT treatment of signal transduction disorders
 PS Example 8; Page 88; 138pp; English.
 CC This is a consensus sequence derived from known CLK
 CC serine/threonine kinases. Degenerate primers based on this and
 CC another consensus peptide (see W49920) were used in the RT-PCR
 CC used as probes to isolate novel mCLK2, mCLK3 and mCLK4 clones
 CC (see W49912-14) from a mouse embryo 11.5 p.c. 12AP cDNA library.
 CC The invention relates to novel proteins (see W49906-14) involved in
 CC cellular signal transduction and to the nucleic acids (see
 CC W17097-99) coding for them, and provides vectors, host cells,
 CC purified recombinant proteins, methods for identifying compounds
 CC that activate or inhibit the novel proteins, and methods for the
 CC diagnosis and treatment of diseases associated with the novel
 CC proteins.
 SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 LXXR 10
 | |
 Db 4 LAAR 7

RESULT 11

W50983
 ID W50983 standard; peptide; 7 AA.
 AC W50983;
 DT 03-AUG-1998 (first entry)
 DE Growth hormone releasing factor peptide fragment intermediate.
 KW Growth hormone releasing factor; GRF; synthesis; intermediate.
 OS Synthetic.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Modified_site 1 /note= "Alpha-amino protecting group"
 FT Modified_site 3 /note= "Side-chain protecting group"
 FT Modified_site 5 /note= "Side-chain protecting group"
 FT Modified_site 6 /note= "Side-chain protecting group"
 FT Modified_site 6 /note= "Side-chain protecting group"
 PN CA1339631-C.
 PD 13-JAN-1998.
 PF 12-OCT-1988; 579916.
 PR 12-OCT-1988; CA-579916.
 PA (BOEH) BOEHRINGER INGELHEIM CANADA LTD.
 PI Gauthier J;
 DR WPI; 98-207809/19.
 PT Preparation of human growth hormone releasing factor (hGRF) fragment
 PT - by assembling blocks of a few amino acids each on separate resins,
 PT cleaving, coupling, and deprotecting; is adaptable to scale-up
 PS Claim 1; Page 22; 25pp; English.
 CC The invention relates to an improved process for preparing an active
 CC fragment of human growth hormone releasing factor (hGRF). The invention
 CC relates to an efficient process for preparing the amidated fragment of
 CC hGRF containing the first 29 amino acids of the N-terminal portion
 CC and comprises assembling blocks of a few amino acids each on separate

CC resins, cleaving, coupling, and deprotecting. The present sequence
 CC represents a specifically claimed intermediate of the process.
 SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 |
 Db 2 LSAR 5

RESULT 12
 W81413
 ID W81413 standard; peptide; 7 AA.
 AC W81413;
 DT 22-JAN-1999 (first entry)
 DE PTK kinase domain conserved region flanking peptide sequence 1.
 KW PTK; receptor; protein tyrosine kinase; recombinant; grafting;
 KW diagnosis; tumour; skin transplant; connective tissue.
 OS Rattus sp.
 PN US5837448-A.
 PD 17-NOV-1998.
 PF 02-MAY-1994; 237401.
 PR 15-MAY-1992; US-884486.
 PR 02-MAY-1994; US-237401.
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 PI Lai CHC. Lemke GE;
 DR WPI; 99-023436/02.
 PT Nucleic acids encoding protein tyrosine kinase subtypes - for
 PT identification of new subtypes and treatment of diseases associated
 PT with the kinase
 PS Example 1; Column 12; 47pp; English.
 CC Sequences W81413 to W81415 represent peptides flanking a highly conserved
 CC region of the kinase domain of receptor protein tyrosine kinase (PTK).
 CC These are used for designing primers to amplify PTK related sequences.
 CC The invention provides sequences V65308 to V65313, V65315, and V65317 to
 CC V65319 that encode proteins having a tyrosine kinase domain and a tissue
 CC expression pattern of a receptor PTK subtype selected from tyro-1,
 CC tyro-2, tyro-3, tyro-4, tyro-5, tyro-6, tyro-8, tyro-10, tyro-11, and
 CC tyro-12, respectively. The polynucleotides are useful for the detection
 CC of tyrosine kinase domain sequences and detection of tissue expression
 CC patterns of PTK subtypes. The cDNAs can also be injected into oocytes,
 CC the protein expressed, and expression products screened for using
 CC antibodies against tyrosine kinase epitopes. These subtypes sequences can
 CC be used for the design of oligonucleotides, for use in amplification
 CC reactions to isolate other subtype sequences. These detection protocols
 CC are used in the diagnosis of diseases associated with (receptor) PTKs.
 CC Recombinant vectors expressing the subtypes can be used to treat related
 CC diseases e.g. tumours, by introduction of the vectors into skin
 CC transplants, then grafting these into the connective tissue of the
 CC dermis, thus specifically targeting tumours as the proteins are released
 CC from the matrix.
 SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 |
 Db 4 LAAR 7

RESULT 13
 Y04013
 ID Y04013 standard; peptide; 7 AA.
 AC Y04013;
 DT 23-JUN-1999 (first entry)
 DE Peptide inhibiting the oncogenic action of p21 ras.

KW p21 ras; adenocarcinoma; colon cancer; pancreatic carcinoma;
 KW neuroblastoma; oncogenesis inhibition; antitumour; anticancer.
 OS Synthetic.
 OS Homo sapiens.
 PN WO9710836-A1.
 PD 27-MAR-1997.
 PF 20-SEP-1996; U15098.
 PR 21-SEP-1995; US-531525.
 PR 21-SEP-1995; US-004091.
 PA (INNA-) INNAPHARMA INC.
 PI Abajlan HB, Hlavka JJ, Kende AS, Noble JF, Pincus MR;
 DR WPI; 97-202613/18.

PT Peptides, cyclised peptides and peptidomimetics - for use in
 PT inhibiting the oncogenic or transforming action of p21 ras protein
 PS Claim 1; Page 93; 100pp; English.
 CC The patent discloses peptides, cyclised peptides and peptidomimetics
 CC capable of inhibiting the oncogenic action of p21 ras. They may be
 CC used in the treatment of adenocarcinomas of the colon, pancreatic
 CC carcinomas, neuroblastomas and other cancers of undefined germ cell
 CC origin. The cyclised peptides correspond to domains of the oncogenic
 CC ras protein which are most flexible and important in interacting with
 CC target proteins upstream and downstream from ras. The peptidomimetics
 CC are obtained by molecular modelling. The present sequence represents
 CC one of the preferred peptides.
 SQ Sequence 7 AA;

Query Match 52.9%; Score 9; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 |
 Db 3 LAAR 6

RESULT 14
 R36108
 ID R36108 standard; protein; 8 AA.
 AC R36108;
 DT 24-MAY-1993 (first entry)
 DE Hepatitis C virus (HCV) epitope Ep36.
 KW Hepatitis; liver disease; HCV; monoclonal antibody; epitope;
 KW immobilised reagent; immunoassay; diagnosis; detection; treatment;
 KW infection.
 OS Hepatitis C virus type 1.
 PN WO9300365-A.
 PD 07-JAN-1993.
 PF 24-JUN-1992; U05388.
 PR 24-JUN-1991; US-722489.
 PA (CHIR) CHIRON CORP.
 PI Chien DY, Rutter W;
 DR WPI; 93-036334/04.
 PT Polypeptide(s) comprising truncated hepatitis C virus sequences -
 PT for detection, prevention and treatment of hepatitis C infection
 PS Example A; Page 38; 80pp; English.
 CC This octamer was found to be immunoreactive with anti-HCV anti-sera.
 CC In the epitope mapping experiment three different samples of anti-sera
 CC were reacted with the peptide octamer, and then incubated with of
 CC HRP-labelled goat anti-human Ig antiser, to enable detection of
 CC binding. This epitope starts from amino acid 2345 of the HCV
 CC polyprotein.
 CC This was found to be a strong epitope.
 SQ Sequence 8 AA;

Query Match 52.9%; Score 9; DB 1; Length 8;
 Best Local Similarity 50.0%; Pred. No. 1.5e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 |
 Db 3 LASR 6

RESULT 15
R36109
ID R36109 standard; protein; 8 AA.
AC R36109; 24-MAY-1993 (first entry)
DT Hepatitis C virus (HCV) epitope.
DE Hepatitis; liver disease; HCV1; monoclonal antibody; epitope;
KW immobilised reagent; immunoassay; diagnosis; detection; treatment;
KW infection.
OS Hepatitis C virus type 1.
PN WO9300365-A.
PD 07-JAN-1993.
PF 24-JUN-1992; U05388.
PR 24-JUN-1991; US-722489.
PA (CHIR) CHIRON CORP.
PI Chien DY, Rutter W;
DR WPI; 93-036334/04.
PT Polypeptide(s) comprising truncated hepatitis C virus sequences -
for detection, prevention and treatment of hepatitis C infection
PS Example A; Page 38; 80pp; English.
CC This octamer was found to be immunoreactive with anti-HCV anti-sera.
CC In the epitope mapping experiment three different samples of anti-sera
CC were reacted with the peptide octamer, and then incubated with
CC HRP-labelled goat anti-human Ig antiserum, to enable detection of
CC binding. This epitope starts from amino acid 2346 of the HCV
CC polyprotein.
SQ Sequence 8 AA;

Query Match 52.9%; Score 9; DB 1; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.5e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 7 LXXR 10
| |
Db 2 LASR 5

Search completed: February 8, 2000, 01:29:34
Job time: 1746 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 11:54:09 ; Search time 117.7 Seconds
(without alignments)
4.008 Million cell updates/sec

Title: US-08-653-294-2
Perfect score: 17
Sequence: 1 XXXXXLXXR 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : PIR_62.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 9 | 52.9 | 8 | I53411 | histone H2A.X - hu |
| 2 | 9 | 52.9 | 19 | E56661 | S-locus specific g |
| 3 | 9 | 52.9 | 19 | S21285 | major integral mem |
| 4 | 9 | 52.9 | 29 | I37301 | MHC class II histo |
| 5 | 9 | 52.9 | 30 | PH0882 | Ig kappa chain V r |
| 6 | 9 | 52.9 | 32 | C28393 | calcietrin - marb |
| 7 | 9 | 52.9 | 33 | T08018 | ycf12 protein - Ch |
| 8 | 9 | 52.9 | 33 | A05010 | hypothetical prote |
| 9 | 9 | 52.9 | 34 | PQ0328 | Aada2 protein - Ps |
| 10 | 9 | 52.9 | 34 | PS0065 | lysis protein t - |
| 11 | 9 | 52.9 | 36 | I84732 | cytotoxic T-lympho |
| 12 | 9 | 52.9 | 37 | I52772 | complement factor |
| 13 | 9 | 52.9 | 37 | T07292 | hypothetical prote |
| 14 | 9 | 52.9 | 39 | PS0063 | lysis protein t - |
| 15 | 9 | 52.9 | 40 | S70494 | ferritin heavy cha |
| 16 | 9 | 52.9 | 40 | S01102 | hypothetical prote |
| 17 | 9 | 52.9 | 41 | B24910 | probable tyrosine |
| 18 | 9 | 52.9 | 42 | T01609 | hypothetical prote |
| 19 | 9 | 52.9 | 42 | I78580 | hemoglobin gamma-G |
| 20 | 9 | 52.9 | 42 | I58221 | hemoglobin gamma-A |
| 21 | 9 | 52.9 | 44 | 1 RHFG | somatoliberin - pi |
| 22 | 9 | 52.9 | 44 | 1 RHBS | somatoliberin - bo |
| 23 | 9 | 52.9 | 45 | JQ0738 | ribosomal protein |
| 24 | 9 | 52.9 | 45 | C41870 | ribosomal protein |
| 25 | 9 | 52.9 | 45 | S43055 | hypothetical prote |
| 26 | 9 | 52.9 | 47 | D72005 | ltua protein - Chl |
| 27 | 9 | 52.9 | 48 | S59076 | hypothetical prote |
| 28 | 9 | 52.9 | 49 | D71570 | hypothetical prote |
| 29 | 9 | 52.9 | 49 | S63682 | signal transducer |
| 30 | 9 | 52.9 | 49 | I52510 | calpain I - rat (f |

31 9 52.9 50 2 S13241
32 9 52.9 51 2 S31007
33 9 52.9 52 2 JQ1261
34 9 52.9 53 2 A42442
35 9 52.9 55 1 VRGP
36 9 52.9 55 2 S18015
37 9 52.9 55 2 S68336
38 9 52.9 56 2 S65881
39 9 52.9 56 2 S66588
40 9 52.9 57 1 W9BPC7
41 9 52.9 57 2 PT0201
42 9 52.9 57 2 PT0202
43 9 52.9 57 2 F63183
44 9 52.9 58 2 D48652
45 9 52.9 59 2 A53879

ALIGNMENTS

RESULT 1
I55411
histone H2A.X - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 28-Feb-1997
C:Accession: I55411
R:Ivanova, V.S.; Hatch, C.L.; Bonner, W.M.
J: Biol. Chem. 269, 24189-24194, 1994
A:Title: Characterization of the human histone H2A.X gene. Comparison of its promoter
A:Reference number: I55411; MUID:95014156
A:Accession: I55411
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-8 <RES>
A:Cross-references: GB:S73863; NID:9765295
C:Genetics:
A:Gene: H2A.X

Query Match 52.9%; Score 9; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
| |
Db 3 LAAR 6

RESULT 2
E56661
S-locus specific glycoprotein (allele S2) - wild cabbage (fragment)
C:Species: Brassica oleracea (wild cabbage)
C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 05-Jan-1996
C:Accession: E56661
R:Gaude, T.; Demoroy, L.; Dumas, C.
Electrophoresis 12, 646-653, 1991
A:Title: Use of a fast protein electrophoretic purification procedure for N-terminal
A:Reference number: A56661; MUID:92090397
A:Accession: E56661
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-19 <GAU>
A:Experimental source: stigma extracts, var. alboaglabra
A:Note: sequence extracted from NCBI backbone (NCBIP.72299)
C:Comment: This glycoprotein, expressed only in stigmas, plays an important role in t
C:Keywords: glycoprotein; polymorphism

Query Match 52.9%; Score 9; DB 2; Length 19;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10

```
Db 4 LSSR 7
| |
RESULT 3
S21285
A:Title: Integral membrane protein, peroxisomal - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
C:Accession: S21285
R:Chen, N.; Crane, D.I.
B:Biochem. J. 283, 605-610, 1992
A:Title: Induction of the major integral membrane protein of mouse liver peroxisomes by
A:Reference number: S21285; MUID:92246895
A:Accession: S21285
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-19 <CHE>
A>Note: 1-Ala, 16-Ser, and 18-Ala were also found

Query Match 52.9%; Score 9; DB 2; Length 19;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 LXXR 10
| |
Db 7 LTR 10
| |

RESULT 4
I37301
MHC class II histocompatibility antigen HLA-DR beta-3 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
C:Accession: I37301; I37302
R:Louis, P.; Eliaou, J.F.; Kerlan-Candon, S.; Pinet, V.; Vincent, R.; Clot, J.
Immunogenetics 38, 21-26, 1993
A:Title: Polymorphism in the regulatory region of HLA-DRB genes correlating with haploty
A:Reference number: I37300; MUID:93216303
A:Accession: I37301
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-29 <RES>
A:Cross-references: EMBL:X65558; NID:g296268; PIDN:CAA46528.1; PID:g296269
A>Note: this allele is designated DRB3*0101
A:Accession: I37302
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-29 <RE2>
A:Cross-references: EMBL:X65559; NID:g296270; PIDN:CAA46529.1; PID:g296271
A>Note: this allele is designated DRB3*0201
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 52.9%; Score 9; DB 2; Length 29;
Best Local Similarity 50.0%; Pred. No. 5.6e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 LXXR 10
| |
Db 22 LSSR 25
| |

RESULT 5
PH0882
Ig kappa chain V region (REN) - human (fragment)
N:Alternate names: myeloma protein
C:Species: Homo sapiens (man)
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
C:Accession: PH0882
R:Manheimer-Lory, A.; Katz, J.B.; Pillinger, M.; Grossein, C.; Smith, A.; Diamond, B.
J. Exp. Med. 174, 1639-1652, 1991

A:Title: Molecular characteristics of antibodies bearing an anti-DNA-associated idiot
A:Reference number: PH0862; MUID:92078875
A:Accession: PH0882
A:Molecule type: protein
A:Residues: 1-30 <MAN>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:1-23/Region: framework 1
F:24-30/Region: complementarity-determining 1

Query Match 52.9%; Score 9; DB 2; Length 30;
Best Local Similarity 50.0%; Pred. No. 5.8e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 LXXR 10
| |
Db 21 LSSR 24
| |

RESULT 6
C26393
calelectrin - marbled electric ray (fragments)
C:Species: Torpedo marmorata (marbled electric ray)
C:Date: 02-Jun-1988 #sequence_revision 02-Jun-1988 #text_change 18-Jun-1993
C:Accession: C26393
R:Geisow, M.J.; Fritzsche, U.; Hexham, J.M.; Dash, B.; Johnson, T.
Nature 320, 636-638, 1986
A:Title: A consensus amino-acid sequence repeat in Torpedo and mammalian Ca(2)+-depen
A:Reference number: A93379; MUID:86203621
A:Accession: C26393
A:Molecule type: protein
A:Residues: 1-32 <GEI>

Query Match 52.9%; Score 9; DB 2; Length 32;
Best Local Similarity 50.0%; Pred. No. 6.2e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 LXXR 10
| |
Db 13 LSSR 16
| |

RESULT 7
T08018
ycf12 protein - Chlamydomonas reinhardtii chloroplast
C:Species: chloroplast Chlamydomonas reinhardtii
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 21-May-1999
C:Accession: T08018
R:Khrebukova, I.; Spreitzer, R.J.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z16296
A:Accession: T08018
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-33 <KHE>
A:Cross-references: EMBL:U40346; NID:gl101912; PID:gl101913
C:Genetics:
A:Gene: ycf12
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 52.9%; Score 9; DB 2; Length 33;
Best Local Similarity 50.0%; Pred. No. 6.4e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 LXXR 10
| |
Db 26 LSSR 29
| |
```

RESULT 8

A05010
hypothetical protein 33 - liverwort (Marchantia polymorpha) chloroplast
C:Species: chloroplast Marchantia polymorpha
C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 31-Oct-1997
C:Accession: S01581; A05010
R:Umesono, K.; Inokuchi, H.; Shiki, Y.; Takeuchi, M.; Chang, Z.; Fukuzawa, H.; Kohchi, T.
J. Mol. Biol. 203, 299-331, 1988
A:Title: Structure and organization of Marchantia polymorpha chloroplast genome. II. Gen
A:Reference number: S01567; MUID:89068686
A:Accession: S01581
A:Molecule type: DNA
A:Residues: 1-33 <MEM>
A:Cross-references: EMBL:X04465; NID:g11640; PID:g11656
R:Ohshima, K.; Fukuzawa, H.; Kohchi, T.; Shirai, H.; Sano, T.; Sano, S.; Umesono, K.; Shi
Nature 322, 572-574, 1986
A:Title: Chloroplast gene organization deduced from complete sequence of liverwort March
A:Reference number: A38014
A:Contents: annotation; gene organization, sites, features
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 52.9%; Score 9; DB 2; Length 33;
Best Local Similarity 50.0%; Pred. No. 6.4e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 7 LXXR 10
| |
Db 26 LAAR 29

RESULT 9

PQ0228
aadA2 protein - Pseudomonas sp. plasmid R1033, transposon Tn1696 (fragment)
C:Species: Pseudomonas sp.
C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 30-Sep-1993
C:Accession: PQ0228
R:Stokes, H.W.; Hall, R.M.
Plasmid 26, 10-19, 1991
A:Title: Sequence analysis of the inducible chloramphenicol resistance determinant in th
A:Reference number: JQ1201; MUID:92052679
A:Accession: PQ0228
A:Molecule type: DNA
A:Residues: 1-34 <STO>
C:Genetics:
A:Gene: aadA2
A:Genome: plasmid
C:Superfamily: streptomycin 3'-adenylyltransferase

Query Match 52.9%; Score 9; DB 2; Length 34;
Best Local Similarity 50.0%; Pred. No. 6.6e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 7 LXXR 10
| |
Db 10 LASR 13

RESULT 10

PS0065
lysis protein t - phase M1 (fragment)
C:Species: phase M1
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Dec-1996
C:Accession: PS0065
R:Montag, D.; Riede, I.; Eschbach, M.L.; Degen, M.; Henning, U.
J. Mol. Biol. 196, 165-174, 1987
A:Title: Receptor-recognizing proteins of T-even type bacteriophages. Constant and hyper
A:Reference number: A94692; MUID:86011316
A:Accession: PS0065
A:Molecule type: DNA

A:Residues: 1-34 <MON>
C:Superfamily: phase T4 lysis protein t
C:Keywords: host cell lysis

Query Match 52.9%; Score 9; DB 2; Length 34;
Best Local Similarity 50.0%; Pred. No. 6.6e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 7 LXXR 10
| |
Db 31 LASR 34

RESULT 11

I84732
cytotoxic T-lymphocyte-associated protein 4 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 02-Jul-1996
C:Accession: I84732
R:Harper, K.; Balzano, C.; Rouvier, E.; Mattei, M.G.; Luciani, M.F.; Golstein, P.
J. Immunol. 147, 1037-1044, 1991
A:Title: CTLA-4 and CD28 activated lymphocyte molecules are closely related in both m
A:Reference number: I49584; MUID:91318145
A:Accession: I84732
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-36 <RES>
A:Cross-references: GB:M74363; NID:g181187; PID:g553245

Query Match 52.9%; Score 9; DB 2; Length 36;
Best Local Similarity 50.0%; Pred. No. 7e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 7 LXXR 10
| |
Db 15 LATR 18

RESULT 12

I55272
complement factor H - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
C:Accession: I55272
R:Wik, D.P.; Keeney, J.B.; Munoz-Canoves, P.; Chaplin, D.D.; Tack, B.F.
J. Biol. Chem. 263, 16720-16724, 1988
A:Title: Structure of the murine complement factor H gene.
A:Reference number: I55272; MUID:89034160
A:Accession: I55272
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-37 <RES>
A:Cross-references: GB:M23553; NID:g340946; PID:AAA39530.1; PID:g554202
C:Superfamily: complement factor H; complement factor H repeat homology

Query Match 52.9%; Score 9; DB 2; Length 37;
Best Local Similarity 50.0%; Pred. No. 7.2e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 7 LXXR 10
| |
Db 21 LSAR 24

RESULT 13

T07292
hypothetical protein ycf12 - Chlorella vulgaris chloroplast
C:Species: chloroplast Chlorella vulgaris
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 22-Jun-1999
C:Accession: T07292

R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakas
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chlo
A:Reference number: Z15985; MUID:97303241

A:Accession: T07292
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-37 <WAK>

A:Cross-references: EMBL:AB001684; NID:d1110444; PID:d1021518

C:Genetics:

A:Gene: ycf12

A:Genome: chloroplast

C:Keywords: chloroplast

Query Match 52.9%; Score 9; DB 2; Length 37;
Best Local Similarity 50.0%; Pred. No. 7.2e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10

Db 30 LASR 33

RESULT 14

PS0063

lysis protein t - phage Ox2 (fragment)

C:Species: phage Ox2

C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 20-Sep-1999

C:Accession: PS0063

R:Montag, D.; Riede, I.; Eschbach, M.L.; Degen, M.; Henning, U.

J. Mol. Biol. 196, 165-174, 1987

A:Title: Receptor-recognizing proteins of T-even type bacteriophages. Constant and hyper

A:Reference number: A94692; MUID:88011316

A:Accession: PS0063

A:Molecule type: DNA

A:Residues: 1-39 <MON>

A:Cross-references: NID:gx50575; NID:gx15124; PIDN:CAA29159.1; PID:gx15127

C:Superfamily: phage T4 lysis protein t

C:Keywords: host cell lysis

Query Match 52.9%; Score 9; DB 2; Length 39;
Best Local Similarity 50.0%; Pred. No. 7.6e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10

Db 31 LASR 34

RESULT 15

S70494

ferritin heavy chain - Absidia spinosa (fragment)

C:Species: Absidia spinosa

C>Date: 19-Mar-1998 #sequence_revision 19-Mar-1998 #text_change 07-May-1999

C:Accession: S70494

R:Carriano, C.J.; Boehnke, R.; Matzanke, B.F.

FEBS Lett. 390, 261-264, 1996

A:Title: Fungal ferritins: the ferritin from mycelia of Absidia spinosa is a bacteriofer

A:Reference number: S70494; MUID:96311417

A:Accession: S70494

A:Molecule type: protein

A:Residues: 1-40 <CAR>

A:Experimental source: strain Tue268

C:Superfamily: bacterioferritin

C:Keywords: heme; iron storage

Query Match 52.9%; Score 9; DB 2; Length 40;
Best Local Similarity 50.0%; Pred. No. 7.8e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10

Db 19 LAAR 22

Search completed: February 7, 2000, 11:54:11
Job time: 24321 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 00:59:40 : Search time 63.71 Seconds
(without alignments)
4.688 Million cell updates/sec

Title: US-08-653-294-2
Perfect score: 17
Sequence: 1 XXXXXXLXXR 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 82229 segs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SwissProt_38.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 9 | 52.9 | 33 | 1 YC12_CHLRE | P50370 chlamydomon |
| 2 | 9 | 52.9 | 33 | 1 YC12_MARPO | P31560 marchantia |
| 3 | 9 | 52.9 | 34 | 1 VLXS_BPM1 | P08229 bacterioph |
| 4 | 9 | 52.9 | 37 | 1 YC12_CHLVU | P56328 chlorella v |
| 5 | 9 | 52.9 | 39 | 1 VLXS_BPOX2 | P08230 bacterioph |
| 6 | 9 | 52.9 | 44 | 1 SLIB_BOVIN | P01288 bos taurus |
| 7 | 9 | 52.9 | 44 | 1 SLIB_PIG | P01287 sus scrofa |
| 8 | 9 | 52.9 | 44 | 1 SLIB_SHEEP | P07217 ovis aries |
| 9 | 9 | 52.9 | 45 | 1 RL34_MICLU | P21153 micrococcus |
| 10 | 9 | 52.9 | 45 | 1 RL34_STRCO | P27901 streptomyce |
| 11 | 9 | 52.9 | 45 | 1 SLIB_CYPCA | P42692 cyprinus ca |
| 12 | 9 | 52.9 | 49 | 1 CCMC_RHILV | P45407 rhizobium l |
| 13 | 9 | 52.9 | 51 | 1 VG62_BPM15 | Q05275 mycobacteri |
| 14 | 9 | 52.9 | 52 | 1 V07K_FXMV | P22171 foxtail mos |
| 15 | 9 | 52.9 | 56 | 1 NINE_BP82 | Q37871 bacterioph |
| 16 | 9 | 52.9 | 56 | 1 NINE_ECOLI | Q47270 escherichia |
| 17 | 9 | 52.9 | 57 | 1 V193_BPT7 | P03790 bacterioph |
| 18 | 9 | 52.9 | 59 | 1 SASO_GLOPE | P41371 clostridium |
| 19 | 9 | 52.9 | 59 | 1 SAS2_GLOPE | P21887 clostridium |
| 20 | 9 | 52.9 | 60 | 1 HBXA_ECHGR | P55813 echinococcu |
| 21 | 9 | 52.9 | 61 | 1 DNBI_BFDV | P13893 budgerigar |
| 22 | 9 | 52.9 | 62 | 1 UCRA_BOVIN | P00130 bos taurus |
| 23 | 9 | 52.9 | 62 | 1 VX2A_CVPFS | P18451 porcine tra |
| 24 | 9 | 52.9 | 63 | 1 YORQ_TTV1 | P19301 thermoprote |
| 25 | 9 | 52.9 | 64 | 1 BVCP_NPVBM | P24649 bombyx mori |
| 26 | 9 | 52.9 | 64 | 1 SAS2_GLOBI | P22066 clostridium |
| 27 | 9 | 52.9 | 65 | 1 VLXS_BPPH8 | P05998 bacterioph |
| 28 | 9 | 52.9 | 66 | 1 KED2_ECOLI | P13967 escherichia |
| 29 | 9 | 52.9 | 66 | 1 Y7KD_STRLI | P22401 streptomyce |
| 30 | 9 | 52.9 | 69 | 1 H6_ONCMY | P02315 oncorhynch |
| 31 | 9 | 52.9 | 69 | 1 RL38_ARATH | Q22860 arabidopsis |
| 32 | 9 | 52.9 | 69 | 1 RL38_HUMAN | P23411 homo sapien |
| 33 | 9 | 52.9 | 69 | 1 RL38_LYCES | P46291 lycopersico |
| 34 | 9 | 52.9 | 70 | 1 Y4ON_RHISN | P55599 rhizobium s |

ALIGNMENTS

```

RESULT 1
YC12_CHLRE          71 1 TRY5_ECOLI
ID YC12_CHLRE      72 1 VIP_CAVPO
AC P50370;         72 1 VPI3_BPPH6
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHETICAL 3.3 KD PROTEIN YCF12.
GN YCF12.
OS Chlamydomonas reinhardtii.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2137;
RA KHBREBUKOVA I., SPREITZER R.J.;
RT "Nucleotide sequences of the chloroplast trns-GCU and ycf12 genes of
RT Chlamydomonas reinhardtii."
RL (In) Plant Gene Register PGR95-117.
CC - SIMILARITY: BELONGS TO THE YCF12 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U40346; AAA91171.1;
DR MENDEL: 5177; CHLRE.ycf12.1;
KW Chloroplast; Hypothetical protein.
SQ SEQUENCE 33 AA; 3332 MW; AEC9FF0C CRC32;

```

Query Match 52.9%; Score 9; DB 1; Length 33;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
| |
Db 26 LSAR 29

```

RESULT 2
YC12_MARPO          33 AA.
ID YC12_MARPO      33 AA.
AC P31560;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 01-FEB-1996 (Rel. 33, Last annotation update)
DE HYPOTHETICAL 3.4 KD PROTEIN (ORF 33).
GN YCF12.
OS Marchantia polymorpha (Liverwort).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiopsida;

```

OC Marchantiales; Marchantiaceae; Marchantia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 89068687.
 RA FUKUZAWA H., KOHCHI T., SANO T., SHIRAI H., UMESONO K., INOKUCHI H.,
 RA OZEKI H., OHYAMA K.;
 RT "Structure and organization of Marchantia polymorpha chloroplast
 genome. III. Gene organization of the large single copy region from
 rbcL to trnI(CAU).";
 RL J. Mol. Biol. 203:333-351(1988).
 RN [2]
 RP COMPLETE GENOME.
 RA OHYAMA K., FUKUZAWA H., KOHCHI T., SHIRAI H., SANO T., SANO S.,
 RA UMESONO K., SHIKI Y., TAKEUCHI M., CHANG Z., AOTA S., INOKUCHI H.,
 RA OZEKI H.;
 RT "Chloroplast gene organization deduced from complete sequence of
 liverwort Marchantia polymorpha chloroplast DNA.";
 RL Nature 322:572-574(1986).
 CC -!- SIMILARITY: BELONGS TO THE YCF12 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X04465; CAA28069.1; -.
 DR PIR; A05010; A05010.
 DR PIR; S01581; S01581.
 DR MENDEL; 5293; MARPO:ycf12:1.
 KW Chloroplast; Hypothetical protein.
 SQ SEQUENCE 33 AA; 3386 MW; A80604E5 CRC32;

Query Match 52.9%; Score 9; DB 1; Length 33;
 Best Local Similarity 50.0%; Pred. No. 2.9e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 |
 DB 26 LAAR 29

RESULT 3
 VLVS_BPM1 STANDARD; PRT; 34 AA.
 ID VLVS_BPM1
 AC P08229;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-MAR-1989 (Rel. 10, Last annotation update)
 DE LYSIS PROTEIN (FRAGMENT).
 GN T.
 OS Bacteriophage M1.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 88011316.
 RA MONTAG D., RIEDE I., ESCHBACH M.-L., DEGEN M., HENNING U.;
 RT "Receptor-recognizing proteins of T-even type bacteriophages.
 RT Constant and hypervariable regions and an unusual case of
 RT evolution.";
 RL J. Mol. Biol. 196:165-174(1987).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X05676; CAA29162.1; -.
 DR PIR; PS0065; PS0065.
 KW Phage lysis protein.
 FT NON_TER 34 34
 SQ SEQUENCE 34 AA; 3694 MW; E8C6763A CRC32;

Query Match 52.9%; Score 9; DB 1; Length 34;
 Best Local Similarity 50.0%; Pred. No. 3e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 |
 DB 31 LASR 34

RESULT 4
 YC12_CHLVU STANDARD; PRT; 37 AA.
 ID YC12_CHLVU
 AC P56328;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HYPOTHETICAL 3.9 KD PROTEIN YCF12.
 GN YCF12.
 OS Chlorella vulgaris.
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
 OC Chlorellaceae; Chlorella.
 CC [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=IAM C-27 / TAMIYA;
 RX MEDLINE; 97303241.
 RA WAKASUGI T., NAGAI T., KAPOOR M., SUGITA M., ITO S.,
 RA TSUZUKI J., NAKASHIMA K., TSUZUKI T., SUZUKI Y., HAMADA A., OHTA T.,
 RA INAMURA A., YOSHINAGA K., SUGIURA M.;
 RT "Complete nucleotide sequence of the chloroplast genome from the
 RT green alga Chlorella vulgaris: the existence of genes possibly
 RT involved in chloroplast division.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
 CC -!- SIMILARITY: BELONGS TO THE YCF12 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AB001684; BAA20682.1; -.
 DR PIR; A05010; A05010.
 DR PIR; S01581; S01581.
 DR MENDEL; 5293; MARPO:ycf12:1.
 KW Chloroplast; Hypothetical protein.
 SQ SEQUENCE 37 AA; 3942 MW; 565FE83B CRC32;

Query Match 52.9%; Score 9; DB 1; Length 37;
 Best Local Similarity 50.0%; Pred. No. 3.3e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 |
 DB 30 LASR 33

RESULT 5
 VLVS_BPOX2 STANDARD; PRT; 39 AA.
 ID VLVS_BPOX2
 AC P08230;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-AUG-1988 (Rel. 08, Last annotation update)
 DE LYSIS PROTEIN (FRAGMENT).
 GN T.
 OS Bacteriophage M1.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 88011316.
 RA MONTAG D., RIEDE I., ESCHBACH M.-L., DEGEN M., HENNING U.;
 RT "Receptor-recognizing proteins of T-even type bacteriophages.
 RT Constant and hypervariable regions and an unusual case of
 RT evolution.";
 RL J. Mol. Biol. 196:165-174(1987).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AB001684; BAA20682.1; -.
 DR PIR; A05010; A05010.
 DR PIR; S01581; S01581.
 DR MENDEL; 5293; MARPO:ycf12:1.
 KW Chloroplast; Hypothetical protein.
 SQ SEQUENCE 39 AA; 3942 MW; 565FE83B CRC32;

QY 7 LXXR 10
 |
 DB 30 LASR 33

RESULT 5
 VLVS_BPOX2 STANDARD; PRT; 39 AA.
 ID VLVS_BPOX2
 AC P08230;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-AUG-1988 (Rel. 08, Last annotation update)
 DE LYSIS PROTEIN (FRAGMENT).
 GN T.

OS Bacteriophage OX2.
 OC Viruses; dsDNA viruses, no RNA stage; tailed phages; Myoviridae;
 RN T4-like phages.
 RP [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 88011316.
 RA MONTAG D., RIEDE I., ESCHBACH M.-L., DEGEN M., HENNING U.;
 RT "Receptor-recognizing proteins of T-even type bacteriophages.
 RT Constant and hypervariable regions and an unusual case of
 RT evolution.";
 RL J. Mol. Biol. 196:165-174(1987).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X05675; CAA29159.1; --
 DR PIR; PS0063; PS0063.
 KW Phage lysis protein. 39
 FT NON_TER 39
 FT SEQUENCE 39 AA; 4176 MW; B96B98A6 CRC32;
 SQ SEQUENCE 39 AA; 4176 MW; B96B98A6 CRC32;
 Query Match 52.9%; Score 9; DB 1; Length 39;
 Best Local Similarity 50.0%; Pred. No. 3.5e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 7 LXXR 10
 Db 31 LASR 34
 RESULT 6
 SLIB_BOVIN STANDARD; PRT; 44 AA.
 ID SLIB_BOVIN
 AC P01288;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE SOMATOLIBERIN (GROWTH HORMONE-RELEASING FACTOR) (GRF) (GROWTH
 DE HORMONE-RELEASING HORMONE) (GHRH).
 GN GHRH.
 OS Bos taurus (Bovine), and Capra hircus (Goat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovinae; Bos.
 RN [1]
 RP SEQUENCE.
 RC SPECIES=BOVINE;
 RX MEDLINE; 84127993.
 RA ESCH F., BOHLEN P., LING N., BRAZEAU P., GUILLEMIN R.;
 RT "Isolation and characterization of the bovine hypothalamic growth
 RT hormone releasing factor.";
 RL Biochem. Biophys. Res. Commun. 117:772-779(1983).
 RP [2]
 RP SEQUENCE.
 RC SPECIES=C.HIRCUS;
 RX MEDLINE; 85096956.
 RA BRAZEAU P., BOHLEN P., ESCH F., LING N., WEHREBERG W.B.,
 RA GUILLEMIN R.;
 RT "Growth hormone-releasing factor from ovine and caprine hypothalamus:
 RT isolation, sequence analysis and total synthesis.";
 RL Biochem. Biophys. Res. Commun. 125:606-614(1984).
 CC -1- FUNCTION: GRF IS RELEASED BY THE HYPOTHALAMUS AND ACTS ON THE
 CC ADENOHYPOPHYSSE TO STIMULATE THE SECRETION OF GROWTH HORMONE.
 CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 DR PROSITE; PS00260; GLUCAGON; 1.
 DR PFAM; PF00123; hormone2; 1.
 RP SEQUENCE.

KW Glucagon family; Amidation; Hypothalamus.
 FT MOD_RES 44 44
 SQ SEQUENCE 44 AA; 5109 MW; 9FA0D0EE CRC32;
 Query Match 52.9%; Score 9; DB 1; Length 44;
 Best Local Similarity 50.0%; Pred. No. 4e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 7 LXXR 10
 Db 17 LSAR 20
 RESULT 7
 SLIB_PIG STANDARD; PRT; 44 AA.
 ID SLIB_PIG
 AC P01287;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE SOMATOLIBERIN (GROWTH HORMONE-RELEASING FACTOR) (GRF) (GROWTH
 DE HORMONE-RELEASING HORMONE) (GHRH).
 GN GHRH.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 RN [1]
 RP SEQUENCE.
 RC TISSUE=HYPOTHALAMUS;
 RX MEDLINE; 84079886.
 RA BOHLEN P., ESCH F., BRAZEAU P., LING N., GUILLEMIN R.;
 RT "Isolation and characterization of the porcine hypothalamic growth
 RT hormone releasing factor.";
 RL Biochem. Biophys. Res. Commun. 116:726-734(1983).
 CC -1- FUNCTION: GRF IS RELEASED BY THE HYPOTHALAMUS AND ACTS ON THE
 CC ADENOHYPOPHYSSE TO STIMULATE THE SECRETION OF GROWTH HORMONE.
 CC -1- MISCELLANEOUS: THE CARBOXYL-AMIDATED SOMATOLIBERIN IS TWICE AS
 CC ACTIVE AS THAT HAVING A FREE CARBOXYL END.
 CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 DR PIR; A01553; RHFG.
 DR PROSITE; PS00260; GLUCAGON; 1.
 DR PFAM; PF00123; hormone2; 1.
 KW Glucagon family; Amidation; Hypothalamus.
 FT MOD_RES 44 44
 SQ SEQUENCE 44 AA; 5110 MW; B75EF294 CRC32;
 Query Match 52.9%; Score 9; DB 1; Length 44;
 Best Local Similarity 50.0%; Pred. No. 4e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 7 LXXR 10
 Db 17 LSAR 20
 RESULT 8
 SLIB_SHEEP STANDARD; PRT; 44 AA.
 ID SLIB_SHEEP
 AC P07217;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE SOMATOLIBERIN (GROWTH HORMONE-RELEASING FACTOR) (GRF) (GROWTH
 DE HORMONE-RELEASING HORMONE) (GHRH).
 GN GHRH.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Caprinae; Ovis.
 RN [1]
 RP SEQUENCE.

RX MEDLINE: 85096956.
 RA BRAZEAU P., BOHLEN P., ESCH F., LING N., WEHREBERG W.B.,
 RA GUILLEMIN R.;
 RT "Growth hormone-releasing factor from ovine and caprine hypothalamus:
 RT Isolation, sequence analysis and total synthesis.";
 RL Biochem. Biophys. Res. Commun. 125:606-614(1984).
 CC -!- FUNCTION: GRF IS RELEASED BY THE HYPOTHALAMUS AND ACTS ON THE
 CC ADENOHYPOPHYSSE TO STIMULATE THE SECRETION OF GROWTH HORMONE.
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 DR PROSITE: PS00260; GLUCAGON; 1.
 DR PFAM: PF00123; hormone2; 1.
 DR KW Glucagon family; Amidation; Hypothalamus.
 FT MOD_RES 44 44
 FT SEQUENCE 44 AA: 5123 MW; D97898C6 CRC32;

Query Match 52.9%; Score 9; DB 1; Length 44;
 Best Local Similarity 50.0%; Pred. No. 4e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 DB 17 LSAR 20

RESULT 9

RL34_MICLU
 ID RL34_MICLU STANDARD; PRT; 45 AA.
 AC P21153;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE 50S RIBOSOMAL PROTEIN L34.
 GN RPMH;
 OS Micrococcus luteus (Micrococcus lysodeikticus).
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Micrococccineae; Micrococccaceae; Micrococcus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 91033019.
 RA FUJITA M.Q., YOSHIKAWA H., OGASAWARA N.;
 RT "Structure of the dnaA region of Micrococcus luteus: conservation and
 RT variations among eubacteria.";
 RL Gene 93:73-78(1990).
 CC -!- SIMILARITY: BELONGS TO THE L34P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M34006; AAA25314.1;
 DR PIR: JQ0738; JQ0738.
 DR PROSITE: PS00784; RIBOSOMAL_L34; 1.
 DR PFAM: PF00468; Ribosomal_L34; 1.
 KW Ribosomal protein.
 SQ SEQUENCE 45 AA: 5309 MW; 46A593A0 CRC32;

Query Match 52.9%; Score 9; DB 1; Length 45;
 Best Local Similarity 50.0%; Pred. No. 4.1e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 DB 32 LSAR 35

RESULT 10

RL34_STRCO
 ID RL34_STRCO STANDARD; PRT; 45 AA.
 AC P27901;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-DEC-1992 (Rel. 24, Last annotation update)
 DE 50S RIBOSOMAL PROTEIN L34.
 GN RPMH;
 OS Streptomyces coelicolor.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=A3(2);
 RC MEDLINE: 92250416.
 RA CALCUTT M.J., SCHMIDT F.J.;
 RT "Conserved gene arrangement in the origin region of the Streptomyces
 RT coelicolor chromosome.";
 RL J. Bacteriol. 174:3220-3226(1992).
 CC -!- SIMILARITY: BELONGS TO THE L34P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M82836; AAA26733.1;
 DR PROSITE: PS00784; RIBOSOMAL_L34; 1.
 DR PFAM: PF00468; Ribosomal_L34; 1.
 KW Ribosomal protein.
 SQ SEQUENCE 45 AA: 5296 MW; 61C46B4B CRC32;

Query Match 52.9%; Score 9; DB 1; Length 45;
 Best Local Similarity 50.0%; Pred. No. 4.1e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 DB 32 LATR 35

RESULT 11

SLIB_CYPCA
 ID SLIB_CYPCA STANDARD; PRT; 45 AA.
 AC P42692;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE SOMATOLIBERIN (GROWTH HORMONE-RELEASING FACTOR) (GRF) (GROWTH
 DE HORMONE-RELEASING HORMONE) (GHRH).
 OS Cyprinus carpio (Common carp).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 CC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
 CC Cyprinidae; Cyprininae; Cyprinidae; Cyprinidae; Cyprinus.
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RC TISSUE=HYPOTHALAMUS;
 RX MEDLINE: 93116845.
 RA VAUGHAN J.M., RIVIER J., SPIESS J., PENG C., CHANG J.P., PETER R.E.,
 RA VALE W.;
 RT "Isolation and characterization of hypothalamic growth-hormone
 RT releasing factor from common carp, *Cyprinus carpio*.";
 RL Neuroendocrinology 56:539-549(1992).
 CC -!- FUNCTION: GRF IS RELEASED BY THE HYPOTHALAMUS AND ACTS ON THE
 CC ADENOHYPOPHYSSE TO STIMULATE THE SECRETION OF GROWTH HORMONE.
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 DR PROSITE: PS00260; GLUCAGON; 1.
 DR PFAM: PF00123; hormone2; 1.
 KW Glucagon family; Hypothalamus.
 SQ SEQUENCE 45 AA: 4979 MW; 51F6B952 CRC32;

```
Query Match 52.9%; Score 9; DB 1; Length 45;
Best Local Similarity 50.0%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
DB 17 LSAR 20

RESULT 12
CCMC_RHLV STANDARD; PRT; 49 AA.
AC P45407;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE HEME EXPORTER PROTEIN C (CYTOCHROME C-TYPE BIOGENESIS PROTEIN CYCZ)
DE (FRAGMENT).
OS Rhizobium leguminosarum (biovar viciae).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8401;
RX MEDLINE; 94292432.
RA VARGAS C., WU G., DAVIES A., DOWNIE J.A.;
RT "Identification of a gene encoding a thioredoxin-like product
RT necessary for cytochrome c biosynthesis and symbiotic nitrogen
RT fixation in Rhizobium leguminosarum.";
RL J. Bacteriol. 176:4117-4123(1994).
CC -1- FUNCTION: REQUIRED FOR THE EXPORT OF HEME TO THE PERIPLASM FOR THE
CC BIOGENESIS OF C-TYPE CYTOCHROMES.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE CMC/CYCZ/HELIC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X79307; CAA55886.1; --
KW Cytochrome c-type biogenesis; Transport; Transmembrane;
FT NON_TER 1
SQ SEQUENCE 49 AA; 5966 MW; C09F8BA6 CRC32;

Query Match 52.9%; Score 9; DB 1; Length 49;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
DB 40 LAAR 43

RESULT 13
VG62_BPML5 STANDARD; PRT; 51 AA.
ID VG62_BPML5
AC Q05275;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE GENE 62 PROTEIN (GP62).
GN 62.
OS Mycobacteriophage L5.
OC Viruses.

[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE; 93211282.
RA HATFULL G.F., SARKIS G.J.;
RT "DNA sequence, structure and gene expression of mycobacteriophage L5:
RT a phage system for mycobacterial genetics.";
RL Mol. Microbiol. 7:395-405(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z18946; CAA79438.1; --
DR PIR; S31007; S31007.
SQ SEQUENCE 51 AA; 5883 MW; DE68F9AC CRC32;

Query Match 52.9%; Score 9; DB 1; Length 51;
Best Local Similarity 50.0%; Pred. No. 4.7e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
DB 28 LATR 31

RESULT 14
V07K_FXMV STANDARD; PRT; 52 AA.
ID V07K_FXMV
AC P22171;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-MAY-1992 (Rel. 22, Last annotation update)
DE 7 KD PROTEIN (ORF 4).
OS Foxtail mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91374015.
RA BANCROFT J.B., ROULEAU M., JOHNSTON R., PRINS L., MACKIE G.A.;
RT "The entire nucleotide sequence of foxtail mosaic virus RNA.";
RL J. Gen. Virol. 72:2173-2181(1991).
CC -1- SIMILARITY: TO OTHER 7 KD PROTEINS (ORF4) FROM POTEXVIRUSES
CC AND CARLAVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62730; AAA43829.1; --
DR PIR; JQ1261; JQ1261.
KW Transmembrane.
SQ SEQUENCE 52 AA; 5849 MW; 1367F80D CRC32;

Query Match 52.9%; Score 9; DB 1; Length 52;
Best Local Similarity 50.0%; Pred. No. 4.8e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
DB 20 LSTR 23

RESULT 15
```

```

NINE BP82
ID NINE_BP82 STANDARD; PRT; 56 AA.
AC Q37871;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE NINE PROTEIN.
GN NINE.
OS Bacteriophage 82.
OC Viruses; GSDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
OC Lambda phage group.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96196428.
RA MAHDI A.A., SHARPLES G.J., MANDAL T.N., LLOYD R.G.;
RT "Holliday junction resolvases encoded by homologous rusa genes in
RL J. Mol. Biol. 257:561-573(1996).
CC -!- SIMILARITY: TO LAMBDA NINE (NIN60) AND STRONG, TO E.COLI NINE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X92588; CAA63328.1; .. D1CD97A0 CRC32;
SQ SEQUENCE 56 AA; 6493 MW;

```

Query Match 52.9%; Score 9; DB 1; Length 56;
Best Local Similarity 50.0%; Pred. No. 5.2e+02;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 7 LXXR 10
   | |
Db 49 LAAR 52

```

Search completed: February 8, 2000, 00:59:42
Job time: 3771 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 13:17:25 ; Search time 209.03 Seconds
(without alignments)
3.317 Million cell updates/sec

Title: US-08-653-294-2
Perfect score: 17
Sequence: 1 XXXXXLXXR 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

SPTREMBL12:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 9 | 52.9 | 8 | 015896 | 015896 babesia bov |
| 2 | 9 | 52.9 | 14 | 052840 | 052840 rhizobium l |
| 3 | 9 | 52.9 | 17 | 065345 | 065345 gossypium h |
| 4 | 9 | 52.9 | 20 | 051949 | 051949 chlamydia t |
| 5 | 9 | 52.9 | 20 | 051951 | 051951 chlamydia t |
| 6 | 9 | 52.9 | 20 | 069176 | 069176 chlamydia t |
| 7 | 9 | 52.9 | 20 | 053520 | 053520 chlamydia t |
| 8 | 9 | 52.9 | 20 | 046416 | 046416 chlamydia t |
| 9 | 9 | 52.9 | 20 | 085509 | 085509 chlamydia t |
| 10 | 9 | 52.9 | 20 | 085510 | 085510 chlamydia t |
| 11 | 9 | 52.9 | 20 | 085512 | 085512 chlamydia t |
| 12 | 9 | 52.9 | 20 | 085514 | 085514 chlamydia t |
| 13 | 9 | 52.9 | 20 | 085517 | 085517 chlamydia t |
| 14 | 9 | 52.9 | 20 | 085519 | 085519 chlamydia t |
| 15 | 9 | 52.9 | 20 | 085520 | 085520 chlamydia t |
| 16 | 9 | 52.9 | 20 | 085521 | 085521 chlamydia t |
| 17 | 9 | 52.9 | 20 | 085522 | 085522 chlamydia t |
| 18 | 9 | 52.9 | 20 | 085523 | 085523 chlamydia t |
| 19 | 9 | 52.9 | 20 | 085525 | 085525 chlamydia t |
| 20 | 9 | 52.9 | 20 | 085526 | 085526 chlamydia t |

| | | | | | | |
|----|---|------|----|----|--------|--------------------|
| 21 | 9 | 52.9 | 20 | 2 | 085528 | 085528 chlamydia t |
| 22 | 9 | 52.9 | 20 | 2 | 085530 | 085530 chlamydia t |
| 23 | 9 | 52.9 | 20 | 2 | 085531 | 085531 chlamydia t |
| 24 | 9 | 52.9 | 20 | 2 | 085533 | 085533 chlamydia t |
| 25 | 9 | 52.9 | 22 | 4 | Q14825 | Q14825 homo sapien |
| 26 | 9 | 52.9 | 24 | 5 | Q21397 | Q21397 caenorhabdi |
| 27 | 9 | 52.9 | 25 | 4 | Q14402 | Q14402 homo sapien |
| 28 | 9 | 52.9 | 26 | 4 | Q16265 | Q16265 homo sapien |
| 29 | 9 | 52.9 | 26 | 4 | Q16266 | Q16266 homo sapien |
| 30 | 9 | 52.9 | 26 | 5 | Q17251 | Q17251 berce ovata |
| 31 | 9 | 52.9 | 26 | 11 | Q35630 | Q35630 mus musculu |
| 32 | 9 | 52.9 | 27 | 12 | Q92078 | Q92078 hepatitis c |
| 33 | 9 | 52.9 | 27 | 12 | Q92077 | Q92077 hepatitis c |
| 34 | 9 | 52.9 | 27 | 12 | Q92075 | Q92075 hepatitis c |
| 35 | 9 | 52.9 | 27 | 12 | Q92073 | Q92073 hepatitis c |
| 36 | 9 | 52.9 | 28 | 2 | Q47352 | Q47352 escherichia |
| 37 | 9 | 52.9 | 28 | 2 | Q47356 | Q47356 escherichia |
| 38 | 9 | 52.9 | 28 | 2 | Q47357 | Q47357 escherichia |
| 39 | 9 | 52.9 | 28 | 6 | P79406 | P79406 sus scrofa |
| 40 | 9 | 52.9 | 28 | 6 | Q9XS89 | Q9XS89 equus cabal |
| 41 | 9 | 52.9 | 28 | 7 | Q19732 | Q19732 homo sapien |
| 42 | 9 | 52.9 | 29 | 2 | P76833 | P76833 escherichia |
| 43 | 9 | 52.9 | 30 | 2 | Q47355 | Q47355 escherichia |
| 44 | 9 | 52.9 | 31 | 2 | Q47353 | Q47353 escherichia |
| 45 | 9 | 52.9 | 31 | 2 | Q92516 | Q92516 streptomyce |

ALIGNMENTS

RESULT 1
ID 015896 PRELIMINARY; PRT; 8 AA.
AC 015896;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE 01-AUG-1998 (TREMBLrel. 07, Last annotation update)
DE 12D3 ANTIGEN (FRAGMENT).
GN 12D3.
OS Babesia bovis.
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Babesiidae; Babesia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SAFORD ATTENUATED;
RA SILINS G.U., BLAKELEY R.L., RIDDLES P.W.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=SAFORD ATTENUATED;
RA SILINS G.U., BLAKELEY R.L., RIDDLES P.W.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U44917; AAB66362.1;
FT NON_TER 8
SQ SEQUENCE 8 AA: 984 MW: C4347049 CRC32;

Query Match 52.9%; Score 9; DB 5; Length 8;
Best Local Similarity 50.0%; Pred. No. 2.3e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
|
|
Db 2 LATR 5

RESULT 2
Q52840 PRELIMINARY; PRT; 14 AA.
ID Q52840;
AC Q52840;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE HOMOLOGUE WITH C-TERMINUS OF OTHER RHIZOBIUM NODB GENES.

OS Rhizobium loti.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 RN Phyllobacteriaceae; Mesorhizobium.
 RP [1]
 RX SEQUENCE FROM N.A.
 RX MEDLINE; 97002748.
 RA SCOTT D.B., YOUNG C.A., COLLINS-EMERSON J.M., TERZAGHI E.A.,
 RA ROCKMAN E.S., LEWIS P.E., PANKHURST C.E.;
 RT "Novel and complex chromosomal arrangement of Rhizobium loti
 RT nodulation genes";
 RL Mol. Plant Microbe Interact. 9:187-197(1996).
 DR EMBL; L06241; AAB47352.1; -
 SQ SEQUENCE 14 AA; 1600 MW; 30AC228C CRC32;

Query Match 52.9%; Score 9; DB 2; Length 14;
 Best Local Similarity 50.0%; Pred. No. 7.4e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 | |
 Db 6 LSAR 9

RESULT 3
 O65345
 ID O65345 PRELIMINARY; PRT; 17 AA.
 AC O65345;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)
 DE VACUOLAR PROTON ATPASE 16 KDA PROTEOLIPID SUBUNIT (FRAGMENT).
 GN VAPASE.
 OS Gossypium hirsutum (Upland cotton).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA YAMAMOTO E., BAIRD W.V.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF064202; AAC16556.1; -
 FT NON_TER 1
 SQ SEQUENCE 17 AA; 1770 MW; 988B605C CRC32;

Query Match 52.9%; Score 9; DB 10; Length 17;
 Best Local Similarity 50.0%; Pred. No. 8.9e+02;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 | |
 Db 7 LSSR 10

RESULT 4
 O51949
 ID O51949 PRELIMINARY; PRT; 20 AA.
 AC O51949;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-JUN-1998 (TREMBlrel. 06, Last annotation update)
 DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA BOBO L., NOVAK N.G.;
 RL J. Infect. Dis. 0:0-0(1998).
 DR EMBL; AF015547; AAB95375.1; -
 FT NON_TER 1
 SQ SEQUENCE 20 AA; 1971 MW; E5211272 CRC32;

QY 7 LXXR 10
 | |
 Db 16 LPAR 19

Query Match 52.9%; Score 9; DB 2; Length 20;
 Best Local Similarity 50.0%; Pred. No. 1e+03;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 7
 Q53520
 ID Q53520 PRELIMINARY; PRT; 20 AA.

Query Match 52.9%; Score 9; DB 2; Length 20;
 Best Local Similarity 50.0%; Pred. No. 1e+03;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 | |
 Db 16 LPAR 19

RESULT 5
 O51951
 ID O51951 PRELIMINARY; PRT; 20 AA.
 AC O51951;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-JUN-1998 (TREMBlrel. 06, Last annotation update)
 DE OUTER MEMBRANE PROTEIN (FRAGMENT).
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA BOBO L., NOVAK N.G.;
 RL J. Infect. Dis. 0:0-0(1998).
 DR EMBL; AF015549; AAB95377.1; -
 FT NON_TER 1
 SQ SEQUENCE 20 AA; 1971 MW; E5211272 CRC32;

Query Match 52.9%; Score 9; DB 2; Length 20;
 Best Local Similarity 50.0%; Pred. No. 1e+03;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 | |
 Db 16 LPAR 19

RESULT 6
 O69176
 ID O69176 PRELIMINARY; PRT; 20 AA.
 AC O69176;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)
 DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 GN OMP-1.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA HSIEH Y.-H., BOBO L.D.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF065436; AAC17176.1; -
 FT NON_TER 1
 SQ SEQUENCE 20 AA; 1929 MW; 10DF99C2 CRC32;

Query Match 52.9%; Score 9; DB 2; Length 20;
 Best Local Similarity 50.0%; Pred. No. 1e+03;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
 | |
 Db 16 LPAR 19

RESULT 7
 Q53520
 ID Q53520 PRELIMINARY; PRT; 20 AA.

Q53520:
 AC 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
 DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 GN OMP1.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 95318541.
 RA HAYES L.J., PECHARATANA S., BAILEY R.L., HAMPTON T.J., PICKETT M.A.,
 RA MAREY D.C., WATT P.J., WARD M.E.;
 RT "Extent and kinetics of genetic change in the omp1 gene of Chlamydia
 RT trachomatis in two villages with endemic trachoma."
 RL J. Infect. Dis. 172:268-272(1995).
 DR EMBL; S77980; AAB34695.1; -.
 FT NON_TER 1
 SQ SEQUENCE 20 AA; 1959 MW; 8CA17A4A CRC32;
 Query Match 52.9%; Score 9; DB 2; Length 20;
 Best Local Similarity 50.0%; Pred. No. 1e+03;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 7 LXXR 10
 Db 16 LTAR 19
 RESULT 8
 Q46416 ID Q46416 PRELIMINARY; PRT; 20 AA.
 AC Q46416;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
 DE MAJOR OUTER MEMBRANE PROTEIN VARIABLE DOMAIN I (FRAGMENT).
 GN OMP1.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 89173295.
 RA YUAN Y., ZHANG Y.X., WATKINS N.G., CALDWELL H.D.;
 RT "Nucleotide and deduced amino acid sequences for the four variable
 RT domains of the major outer membrane proteins of the 15 Chlamydia
 RT trachomatis serovars."
 RL Infect. Immun. 57:1040-1049(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA LAMPE M.F., SUCHLAND R.J., STAMM W.E.;
 RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
 DR EMBL; L03750; AAB59032.1; -.
 FT NON_TER 1
 FT VARIANT 11 1 V -> A (IN REF. 2 AND 1).
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2017 MW; FB2B029E CRC32;
 Query Match 52.9%; Score 9; DB 2; Length 20;
 Best Local Similarity 50.0%; Pred. No. 1e+03;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 7 LXXR 10
 Db 16 LTAR 19
 RESULT 9
 O85509 ID O85509 PRELIMINARY; PRT; 20 AA.
 AC O85509;

DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
 DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 GN OMP-1.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B;
 RA HSIEH Y.-H., BOBO L.D.;
 RT "Diversity of major outer membrane protein (omp-1) of Chlamydia
 RT trachomatis in trachoma endemic villages, Kongwa, Tanzania."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF070235; AAC25205.1; -.
 FT NON_TER 1
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 1929 MW; 10DF99C2 CRC32;
 Query Match 52.9%; Score 9; DB 2; Length 20;
 Best Local Similarity 50.0%; Pred. No. 1e+03;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 7 LXXR 10
 Db 16 LTAR 19
 RESULT 10
 O85510 ID O85510 PRELIMINARY; PRT; 20 AA.
 AC O85510;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
 DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 GN OMP-1.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BIA;
 RA HSIEH Y.-H., BOBO L.D.;
 RT "Diversity of major outer membrane protein (omp-1) of Chlamydia
 RT trachomatis in trachoma endemic villages, Kongwa, Tanzania."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF070239; AAC25209.1; -.
 FT NON_TER 1
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 1957 MW; DDDE3F15 CRC32;
 Query Match 52.9%; Score 9; DB 2; Length 20;
 Best Local Similarity 50.0%; Pred. No. 1e+03;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 7 LXXR 10
 Db 16 LTAR 19
 RESULT 11
 O85512 ID O85512 PRELIMINARY; PRT; 20 AA.
 AC O85512;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
 DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 GN OMP-1.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B2A;
RA Hsieh Y.-H., BOBO L.D.;
RT "Diversity of major outer membrane protein (omp-1) of Chlamydia
RT trachomatis in trachoma endemic villages, Kongwa, Tanzania.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070251; AAC25221.1; -
FT NON_TER 1 1
FT SEQUENCE 20 AA; 1929 MW; 10DF99C2 CRC32;
SQ

Query Match 52.9%; Score 9; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
| |
Db 16 LTAR 19

RESULT 12
O85514 PRELIMINARY; PRT; 20 AA.
AC O85514;
DT 01-NOV-1998 (TREMREL. 08, Created)
DT 01-NOV-1998 (TREMREL. 08, Last sequence update)
DT 01-NOV-1998 (TREMREL. 08, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
GN OMP-1.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B2A;
RA Hsieh Y.-H., BOBO L.D.;
RT "Diversity of major outer membrane protein (omp-1) of Chlamydia
RT trachomatis in trachoma endemic villages, Kongwa, Tanzania.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070255; AAC25225.1; -
FT NON_TER 1 1
FT SEQUENCE 20 AA; 1929 MW; 10DF99C2 CRC32;
SQ

Query Match 52.9%; Score 9; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
| |
Db 16 LTAR 19

RESULT 13
O85517 PRELIMINARY; PRT; 20 AA.
AC O85517;
DT 01-NOV-1998 (TREMREL. 08, Created)
DT 01-NOV-1998 (TREMREL. 08, Last sequence update)
DT 01-NOV-1998 (TREMREL. 08, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
GN OMP-1.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B2B;
RA Hsieh Y.-H., BOBO L.D.;
RT "Diversity of major outer membrane protein (omp-1) of Chlamydia
RT trachomatis in trachoma endemic villages, Kongwa, Tanzania.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF070259; AAC25229.1; -
FT NON_TER 1 1
FT SEQUENCE 20 AA; 1929 MW; 10DF99C2 CRC32;
SQ

Query Match 52.9%; Score 9; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
| |
Db 16 LTAR 19

RESULT 14
O85519 PRELIMINARY; PRT; 20 AA.
AC O85519;
DT 01-NOV-1998 (TREMREL. 08, Created)
DT 01-NOV-1998 (TREMREL. 08, Last sequence update)
DT 01-NOV-1998 (TREMREL. 08, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
GN OMP-1.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B2C;
RA Hsieh Y.-H., BOBO L.D.;
RT "Diversity of major outer membrane protein (omp-1) of Chlamydia
RT trachomatis in trachoma endemic villages, Kongwa, Tanzania.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070263; AAC25233.1; -
FT NON_TER 1 1
FT SEQUENCE 20 AA; 1929 MW; 10DF99C2 CRC32;
SQ

Query Match 52.9%; Score 9; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LXXR 10
| |
Db 16 LTAR 19

RESULT 15
O85520 PRELIMINARY; PRT; 20 AA.
AC O85520;
DT 01-NOV-1998 (TREMREL. 08, Created)
DT 01-NOV-1998 (TREMREL. 08, Last sequence update)
DT 01-NOV-1998 (TREMREL. 08, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
GN OMP-1.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B2D;
RA Hsieh Y.-H., BOBO L.D.;
RT "Diversity of major outer membrane protein (omp-1) of Chlamydia
RT trachomatis in trachoma endemic villages, Kongwa, Tanzania.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070267; AAC25237.1; -
FT NON_TER 1 1
FT SEQUENCE 20 AA; 1929 MW; 10DF99C2 CRC32;
SQ

Query Match 52.9%; Score 9; DB 2; Length 20;

Best Local Similarity 50.0%; Pred. NO. le+03;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 LXXR 10
| |
Db 16 LTAR 19

Search completed: February 8, 2000, 13:17:26
Job time: 32475 sec

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-2 to: GenEmbl:* out_format : pfs

Date: Feb 8, 2000 4:36 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=Cgml_1\USPTO.spool\US08653294\runat_04022000_160701_15779/app_query.fasta.1
-DB=GenEmbl -QPMF=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEX=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294
-NCFU=6 -ICFU=3 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-2
Query length: 10
Database: GenEmbl.*
Database sequences: 821193
Database length: -1518192014
Search time (sec): 11370.480000

score_list:

| Sequence | Strd | Orig | ZScore | Escore | Len | Documentation |
|-----------------|------|------|--------|---------|-----|-----------------------------------|
| gb_pat:A06946 | + | 9.00 | 90.81 | 4.8e+03 | 14 | A06946 Oligonucleotide CRN2. 10/ |
| gb_pat:A06950 | + | 9.00 | 90.81 | 4.8e+03 | 14 | A06950 Oligonucleotide CRN2. 10/ |
| gb_pat:A06953 | + | 9.00 | 90.81 | 4.8e+03 | 14 | A06953 Oligonucleotide CRN6C. 10/ |
| gb_pat:A06934 | + | 9.00 | 90.81 | 4.8e+03 | 14 | A06934 Oligonucleotide CRN5A. 10/ |
| gb_pat:A06936 | + | 9.00 | 90.81 | 4.8e+03 | 14 | A06936 Oligonucleotide CRN5A. 10/ |
| gb_pat:A06959 | + | 9.00 | 90.81 | 4.8e+03 | 14 | A06959 Oligonucleotide CRN5F. 10/ |
| gb_pat:A06963 | + | 9.00 | 90.81 | 4.8e+03 | 14 | A06963 Oligonucleotide CRN4D. 10/ |
| gb_pat:A06966 | + | 9.00 | 90.81 | 4.8e+03 | 14 | A06966 Oligonucleotide CRN3A. 10/ |
| gb_pat:A06967 | + | 9.00 | 90.81 | 4.8e+03 | 14 | A06967 Oligonucleotide CRN3B. 10/ |
| gb_pat:A06968 | + | 9.00 | 90.81 | 4.8e+03 | 14 | A06968 Oligonucleotide CRN3C. 10/ |
| gb_pat:A06969 | + | 9.00 | 90.81 | 4.8e+03 | 14 | A06969 Oligonucleotide CRN3D. 10/ |
| gb_pat:A06970 | + | 9.00 | 90.81 | 4.8e+03 | 14 | A06970 Oligonucleotide CRN3E. 10/ |
| gb_pat:A06971 | + | 9.00 | 90.81 | 4.8e+03 | 14 | A06971 Oligonucleotide CRN3F. 10/ |
| gb_pat:A42493 | - | 9.00 | 90.81 | 4.8e+03 | 14 | A42493 Sequence 9 from Patent WO |
| gb_pat:A42545 | - | 9.00 | 90.81 | 4.8e+03 | 14 | A42545 Sequence 61 from Patent WO |
| gb_pat:A09096 | - | 9.00 | 90.81 | 4.8e+03 | 14 | I80906 Sequence 23 from patent U |
| gb_pat:A35623 | + | 9.00 | 90.36 | 5.1e+03 | 15 | A35623 Synthetic human IFN-alpha |
| gb_pat:AR055806 | - | 9.00 | 90.36 | 5.1e+03 | 15 | AR055806 Sequence 10 from patent |
| gb_pat:AR056041 | + | 9.00 | 90.36 | 5.1e+03 | 15 | AR056041 Sequence 245 from patent |
| gb_pat:AR056294 | + | 9.00 | 90.36 | 5.1e+03 | 15 | AR056294 Sequence 498 from patent |
| gb_pat:AR056374 | + | 9.00 | 90.36 | 5.1e+03 | 15 | AR056374 Sequence 578 from patent |
| gb_pat:E04104 | - | 9.00 | 90.36 | 5.1e+03 | 15 | E04104 Primer for gene relating |
| gb_pat:E07286 | - | 9.00 | 90.36 | 5.1e+03 | 15 | E07286 Synthetic DNA linkers. 9/ |
| gb_pat:E124585 | + | 9.00 | 90.36 | 5.1e+03 | 15 | E124585 Sequence 13 from patent U |
| gb_pat:I73231 | + | 9.00 | 90.36 | 5.1e+03 | 15 | I73231 Sequence 1 from patent US |
| gb_pat:I76169 | + | 9.00 | 90.36 | 5.1e+03 | 15 | I76169 Sequence 13 from patent U |
| em_pat:E11940 | - | 9.00 | 90.36 | 5.1e+03 | 15 | E11940 Linker. 10/1997 |
| gb_pat:A09270 | - | 9.00 | 89.93 | 5.4e+03 | 16 | A09270 Oligonucleotide (SD). 10/ |
| gb_pat:A10710 | - | 9.00 | 89.93 | 5.4e+03 | 16 | A10710 Oligonucleotide (SD). 12/ |
| gb_pat:A11354 | + | 9.00 | 89.93 | 5.4e+03 | 16 | A11354 Nucleotide sequence 8 fr |
| gb_pat:A11562 | - | 9.00 | 89.93 | 5.4e+03 | 16 | A11562 Oligonucleotide 'SD'. 11/ |
| gb_pat:A15266 | - | 9.00 | 89.93 | 5.4e+03 | 16 | A15266 Oligonucleotide SD. 3/199 |
| gb_pat:A16500 | - | 9.00 | 89.93 | 5.4e+03 | 16 | A16500 Oligonucleotide SD. 3/199 |
| gb_pat:A42492 | - | 9.00 | 89.93 | 5.4e+03 | 16 | A42492 Sequence 8 from Patent WO |
| gb_pat:AR012675 | + | 9.00 | 89.93 | 5.4e+03 | 16 | AR012675 Sequence 14 from patent |
| gb_pat:AR024016 | + | 9.00 | 89.93 | 5.4e+03 | 16 | AR024016 Sequence 4 from patent |
| gb_pat:I24584 | + | 9.00 | 89.93 | 5.4e+03 | 16 | I24584 Sequence 12 from patent U |
| gb_pat:I27931 | - | 9.00 | 89.93 | 5.4e+03 | 16 | I27931 Sequence 103 from patent |
| gb_pat:I72565 | + | 9.00 | 89.93 | 5.4e+03 | 16 | I72565 Sequence 14 from patent U |
| gb_pat:A26513 | - | 9.00 | 89.53 | 5.6e+03 | 17 | A26513 Pt-11S mutant. 4/1995 |
| gb_pat:AR005304 | + | 9.00 | 89.53 | 5.6e+03 | 17 | AR005304 Sequence 34 from patent |
| gb_pat:AR027486 | - | 9.00 | 89.53 | 5.6e+03 | 17 | AR027486 Sequence 11 from patent |

gb_pat:113597 + 9.00 89.53 5.6e+03 17 ! 113597 Sequence 3 from patent
gb_pat:113598 + 9.00 89.53 5.6e+03 17 ! 113598 Sequence 4 from patent
gb_pat:114080 + 9.00 89.53 5.6e+03 17 ! 114080 Sequence 7 from patent

seq_name: gb_pat:A06946

seq_documentation_block:
LOCUS A06946 14 bp DNA PAT 14-OCT-1993
DEFINITION Oligonucleotide CRN2.

ACCESSION A06946

VERSION A06946.1 GI:489032

KEYWORDS synthetic construct.

SOURCE synthetic construct

ORGANISM artificial sequence.

REFERENCE 1 (bases 1 to 14)

AUTHORS Carr,F.J

TITLE Hybridisation probes

JOURNAL Patent: EP 0246864-A 3 25-NOV-1987;

IMPERIAL CHEMICAL INDUSTRIES PLC

LOCATION/Qualifiers

FEATURES

1..14

/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 1 a 4 c 4 g 5 t

ORIGIN

alignment_scores:

Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x A06946 ..

Align seg 1/1 to: A06946 from: 1 to: 14

7 Leu*****Arg 10

|||||||

3 TTGACTGCCCGT 14

seq_name: gb_pat:A06950

seq_documentation_block:
LOCUS A06950 14 bp DNA PAT 14-OCT-1993
DEFINITION Oligonucleotide CRN2.

ACCESSION A06950

VERSION A06950.1 GI:489035

KEYWORDS synthetic construct.

SOURCE synthetic construct

ORGANISM artificial sequence.

REFERENCE 1 (bases 1 to 14)

AUTHORS Carr,F.J

TITLE Hybridisation probes

JOURNAL Patent: EP 0246864-A 8 25-NOV-1987;

IMPERIAL CHEMICAL INDUSTRIES PLC

LOCATION/Qualifiers

FEATURES

1..14

/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 1 a 4 c 4 g 5 t

ORIGIN

alignment_scores:

Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x A06950 ..

Align seg 1/1 to: A06950 from: 1 to: 14

7 Leu*****Arg 10
|||||
3 TTGACTGCCGT 14

seq_name: gb_pat:A06953

seq_documentation_block: 14 bp DNA PAT 14-OCT-1993
LOCUS A06953
DEFINITION Oligonucleotide CRN5C.

ACCESSION A06953

VERSION A06953.1 GI:489038

KEYWORDS synthetic construct.

SOURCE synthetic construct.

ORGANISM artificial sequence.

REFERENCE 1 (bases 1 to 14)

AUTHORS Carr,F.J.

TITLE Hybridisation probes

JOURNAL Patent: EP 0246864-A 11 25-NOV-1987;

IMPERIAL CHEMICAL INDUSTRIES PLC

FEATURES Location/Qualifiers

source 1..14

/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 1 a 3 c 5 g 5 t

ORIGIN

alignment_scores:

Quality: 9.00 Length: 4

Ratio: 2.250 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x A06953 ..

Align seg 1/1 to: A06953 from: 1 to: 14

7 Leu*****Arg 10
|||||
3 TTGACTGCCGT 14

seq_name: gb_pat:A06954

seq_documentation_block: 14 bp DNA PAT 14-OCT-1993
LOCUS A06954
DEFINITION Oligonucleotide CRN5A.

ACCESSION A06954

VERSION A06954.1 GI:489039

KEYWORDS synthetic construct.

SOURCE synthetic construct.

ORGANISM artificial sequence.

REFERENCE 1 (bases 1 to 14)

AUTHORS Carr,F.J.

TITLE Hybridisation probes

JOURNAL Patent: EP 0246864-A 12 25-NOV-1987;

IMPERIAL CHEMICAL INDUSTRIES PLC

FEATURES Location/Qualifiers

source 1..14

/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 0 a 4 c 4 g 6 t

ORIGIN

alignment_scores:

Quality: 9.00 Length: 4

Ratio: 2.250 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x A06954 ..

Align seg 1/1 to: A06954 from: 1 to: 14

7 Leu*****Arg 10
|||||
3 TTGCTGCCGT 14

seq_name: gb_pat:A06956

seq_documentation_block: 14 bp DNA PAT 14-OCT-1993
LOCUS A06956
DEFINITION Oligonucleotide CRN5C.

ACCESSION A06956

VERSION A06956.1 GI:489041

KEYWORDS synthetic construct.

SOURCE synthetic construct.

ORGANISM artificial sequence.

REFERENCE 1 (bases 1 to 14)

AUTHORS Carr,F.J.

TITLE Hybridisation probes

JOURNAL Patent: EP 0246864-A 14 25-NOV-1987;

IMPERIAL CHEMICAL INDUSTRIES PLC

FEATURES Location/Qualifiers

source 1..14

/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 0 a 4 c 5 g 5 t

ORIGIN

alignment_scores:

Quality: 9.00 Length: 4

Ratio: 2.250 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x A06956 ..

Align seg 1/1 to: A06956 from: 1 to: 14

7 Leu*****Arg 10
|||||
3 TTGCTGCCGT 14

seq_name: gb_pat:A06959

seq_documentation_block: 14 bp DNA PAT 14-OCT-1993
LOCUS A06959
DEFINITION Oligonucleotide CRN5F.

ACCESSION A06959

VERSION A06959.1 GI:489044

KEYWORDS synthetic construct.

SOURCE synthetic construct.

ORGANISM artificial sequence.

REFERENCE 1 (bases 1 to 14)

AUTHORS Carr,F.J.

TITLE Hybridisation probes

JOURNAL Patent: EP 0246864-A 17 25-NOV-1987;

IMPERIAL CHEMICAL INDUSTRIES PLC

FEATURES Location/Qualifiers

source 1..14

/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 2 a 4 c 3 g 5 t

ORIGIN

alignment_scores:

Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x A06959 ..

Align seg 1/1 to: A06959 from: 1 to: 14

7 Leu*****Arg 10

|||||

3 TTAGTGTCCCGT 14

seq_name: gb_pat:A06963

seq_documentation_block:

LOCUS A06963 14 bp DNA PAT 14-OCT-1993

DEFINITION Oligonucleotide CRN4D.

ACCESSION A06963

VERSION A06963.1 GI:489048

KEYWORDS

SOURCE

synthetic construct.

artificial sequence.

REFERENCE 1 (bases 1 to 14)

AUTHORS Carr,F.J.

TITLE Hybridisation probes

JOURNAL Patent: EP 0246864-A 21 25-NOV-1987;

IMPERIAL CHEMICAL INDUSTRIES PLC

FEATURES

source

/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 1 a 5 c 4 g 4 t

ORIGIN

alignment_scores:

Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x A06963 ..

Align seg 1/1 to: A06963 from: 1 to: 14

7 Leu*****Arg 10

|||||

3 CTGACTGCCCGT 14

seq_name: gb_pat:A06966

seq_documentation_block:

LOCUS A06966 14 bp DNA PAT 14-OCT-1993

DEFINITION Oligonucleotide CRN3A.

ACCESSION A06966

VERSION A06966.1 GI:489051

KEYWORDS

SOURCE

synthetic construct.

artificial sequence.

REFERENCE 1 (bases 1 to 14)

AUTHORS Carr,F.J.

TITLE Hybridisation probes

JOURNAL Patent: EP 0246864-A 24 25-NOV-1987;

IMPERIAL CHEMICAL INDUSTRIES PLC

FEATURES

source

/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 1 a 4 c 3 g 6 t

ORIGIN

alignment_scores:

Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x A06966 ..

Align seg 1/1 to: A06966 from: 1 to: 14

7 Leu*****Arg 10

|||||

3 TTGACTGCCCGT 14

seq_name: gb_pat:A06967

seq_documentation_block:

LOCUS A06967 14 bp DNA PAT 14-OCT-1993

DEFINITION Oligonucleotide CRN3B.

ACCESSION A06967

VERSION A06967.1 GI:489052

KEYWORDS

SOURCE

synthetic construct.

artificial sequence.

REFERENCE 1 (bases 1 to 14)

AUTHORS Carr,F.J.

TITLE Hybridisation probes

JOURNAL Patent: EP 0246864-A 26 25-NOV-1987;

IMPERIAL CHEMICAL INDUSTRIES PLC

FEATURES

source

/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 1 a 5 c 3 g 5 t

ORIGIN

alignment_scores:

Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x A06967 ..

Align seg 1/1 to: A06967 from: 1 to: 14

7 Leu*****Arg 10

|||||

3 TTGACTGCCCGT 14

seq_name: gb_pat:A06968

seq_documentation_block:

LOCUS A06968 14 bp DNA PAT 14-OCT-1993

DEFINITION Oligonucleotide CRN3C.

ACCESSION A06968

VERSION A06968.1 GI:489053

KEYWORDS

SOURCE

synthetic construct.

artificial sequence.

REFERENCE 1 (bases 1 to 14)

AUTHORS Carr,F.J.

TITLE Hybridisation probes

JOURNAL Patent: EP 0246864-A 27 25-NOV-1987;

IMPERIAL CHEMICAL INDUSTRIES PLC

FEATURES

source

/organism="synthetic construct"

```
FEATURES             Location/Qualifiers
Source               1..14
                    /db_xref="taxon:32630"
BASE COUNT          2 a 4 c 3 g 5 t
ORIGIN

alignment_scores:
  Quality:          9.00      Length:      4
  Ratio:            2.250     Gaps:      0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-08-653-294-2 x A06968 ..
Align seg 1/1 to: A06968 from: 1 to: 14
7 Leu*****Arg 10
|||||
3 TTGACTGCCCGT 14
seq_name: gb_pat:A06969

seq_documentation_block:
LOCUS      A06969      14 bp      DNA      PAT
DEFINITION Oligonucleotide CRN3D.
ACCESSION  A06969
VERSION    A06969.1 GI:489054
KEYWORDS   synthetic construct.
SOURCE     synthetic construct.
ORGANISM   artificial sequence.
REFERENCE  1 (bases 1 to 14)
AUTHORS    Carr,F.J.
TITLE      Hybridisation probes
JOURNAL    Patent: EP 0246864-A 28 25-NOV-1987;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES   Location/Qualifiers
Source     1..14
          /organism="synthetic construct"
          /db_xref="taxon:32630"
BASE COUNT 1 a 5 c 4 g 4 t
ORIGIN

alignment_scores:
  Quality:          9.00      Length:      4
  Ratio:            2.250     Gaps:      0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-08-653-294-2 x A06969 ..
Align seg 1/1 to: A06969 from: 1 to: 14
7 Leu*****Arg 10
|||||
3 TTGACTGCCCGT 14
seq_name: gb_pat:A06970

seq_documentation_block:
LOCUS      A06970      14 bp      DNA      PAT
DEFINITION Oligonucleotide CRN3E.
ACCESSION  A06970
VERSION    A06970.1 GI:489055
KEYWORDS   synthetic construct.
SOURCE     synthetic construct.
ORGANISM   artificial sequence.
REFERENCE  1 (bases 1 to 14)
AUTHORS    Carr,F.J.
TITLE      Hybridisation probes
JOURNAL    Patent: EP 0246864-A 29 25-NOV-1987;
            IMPERIAL CHEMICAL INDUSTRIES PLC
```

```
FEATURES             Location/Qualifiers
Source               1..14
                    /db_xref="taxon:32630"
BASE COUNT          2 a 4 c 4 g 4 t
ORIGIN

alignment_scores:
  Quality:          9.00      Length:      4
  Ratio:            2.250     Gaps:      0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-08-653-294-2 x A06970 ..
Align seg 1/1 to: A06970 from: 1 to: 14
7 Leu*****Arg 10
|||||
3 TTGACTGCCCGT 14
seq_name: gb_pat:A06971

seq_documentation_block:
LOCUS      A06971      14 bp      DNA      PAT
DEFINITION Oligonucleotide CRN3F.
ACCESSION  A06971
VERSION    A06971.1 GI:489056
KEYWORDS   synthetic construct.
SOURCE     synthetic construct.
ORGANISM   artificial sequence.
REFERENCE  1 (bases 1 to 14)
AUTHORS    Carr,F.J.
TITLE      Hybridisation probes
JOURNAL    Patent: EP 0246864-A 30 25-NOV-1987;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES   Location/Qualifiers
Source     1..14
          /organism="synthetic construct"
          /db_xref="taxon:32630"
BASE COUNT 1 a 4 c 5 g 4 t
ORIGIN

alignment_scores:
  Quality:          9.00      Length:      4
  Ratio:            2.250     Gaps:      0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-08-653-294-2 x A06971 ..
Align seg 1/1 to: A06971 from: 1 to: 14
7 Leu*****Arg 10
|||||
3 TTGACTGCCCGT 14
seq_name: gb_pat:A42493

seq_documentation_block:
LOCUS      A42493      14 bp      DNA      PAT
DEFINITION Sequence 9 from Patent WO9502051.
ACCESSION  A42493
VERSION    A42493.1 GI:2297942
KEYWORDS   unidentifed.
SOURCE     unidentifed.
ORGANISM   unclassified.
REFERENCE  1 (bases 1 to 14)
AUTHORS    Schlingensiepen,G., Schlingensiepen,R., Schlingensiepen,K. and
```

Brysch,W.
A PHARMACEUTICAL COMPOSITION COMPRISING ANTISENSE-NUCLEIC ACID FOR
PREVENTION AND/OR TREATMENT OF NEURONAL INJURY, DEGENERATION AND
CELL DEATH AND FOR THE TREATMENT OF NEOPLASMS
JOURNAL Patent: WO 9502051-A 9 19-JAN-1995;
BIOGNOSTIK GES FUER BIOMOLEKUL (DE)
COMMENT Other publication AU 7345694 950206.
FEATURES Location/Qualifiers
source 1..14
/organism="unidentified"
BASE COUNT 3 a 2 c 5 g 4 t
ORIGIN /db_xref="taxon:32644"

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-2 x A42493/rev ..
Align seg 1/1 to reverse of: A42493 from: 1 to: 14

7 Leu*****Arg 10
|||||
13 CTAACCTCACGT 2

seq_name: gb_pat:A42545

seq_documentation_block:
LOCUS A42545 14 bp DNA PAT 06-MAR-1997
DEFINITION Sequence 61 from Patent WO9502051.
ACCESSION A42545
VERSION A42545.1 GI:2297994
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified
REFERENCE 1 (bases 1 to 14)
AUTHORS Schlingensiepen,G., Schlingensiepen,R., Schlingensiepen,K. and
Brysch,W.

TITLE A PHARMACEUTICAL COMPOSITION COMPRISING ANTISENSE-NUCLEIC ACID FOR
PREVENTION AND/OR TREATMENT OF NEURONAL INJURY, DEGENERATION AND
CELL DEATH AND FOR THE TREATMENT OF NEOPLASMS
JOURNAL Patent: WO 9502051-A 61 19-JAN-1995;
BIOGNOSTIK GES FUER BIOMOLEKUL (DE)
COMMENT Other publication AU 7345694 950206.
FEATURES Location/Qualifiers
source 1..14
/organism="unidentified"

BASE COUNT 4 a 1 c 6 g 3 t
ORIGIN /db_xref="taxon:32644"

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-2 x A42545/rev ..
Align seg 1/1 to reverse of: A42545 from: 1 to: 14

7 Leu*****Arg 10
|||||
12 CTCTCTACACGA 1

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-2 to: N_Geneseq_36:* out_format : pfs
Date: Feb 8, 2000 1:27 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+2n.model -DEV=xlp
-O/cgn1_1/USPTO.spool/US08653294/runat_04022000_160701_15807/app_query.fasta.1
-DB=N_Geneseq_36 -OFMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT
-THREADS=1

Search information block:
Query: US-08-653-294-2
Query length: 10
Database: N_Geneseq_36:*
Database sequences: 311585
Database length: 125096042
Search time (sec): 590.520000

| score_list: | Sequence | Strd Orig | ZScore | EScore | Len | Documentation |
|-------------|---------------------|-----------|--------|--------|-----|-----------------------------------|
| | N_Geneseq_36:Q69882 | + | 9.00 | 888.00 | 13 | XbaI/SacI primer. Sequence-direc |
| | N_Geneseq_36:T49563 | + | 9.00 | 888.00 | 13 | XbaI/SacI primer to generate UL9 |
| | N_Geneseq_36:T10946 | - | 9.00 | 96.48 | 13 | Human ribozyme target sequence f |
| | N_Geneseq_36:Q83327 | - | 9.00 | 95.89 | 14 | jub-B antisense oligonucleotide. |
| | N_Geneseq_36:Q83275 | - | 9.00 | 95.77 | 14 | c-jun antisense oligonucleotide. |
| | N_Geneseq_36:T76848 | - | 9.00 | 95.89 | 14 | Probe for HGV NTR detection. Pr |
| | N_Geneseq_36:T48775 | + | 9.00 | 95.89 | 14 | ErbB-2 gene antisense oligonucle |
| | N_Geneseq_36:V48580 | + | 9.00 | 95.89 | 14 | junB gene antisense oligonucleot |
| | N_Geneseq_36:Q35256 | + | 9.00 | 95.35 | 15 | Plant transforming principle dete |
| | N_Geneseq_36:T41816 | + | 9.00 | 95.35 | 15 | HLA allele, HLA-DRB1*08, *12 and |
| | N_Geneseq_36:T49643 | + | 9.00 | 95.35 | 15 | Human CERP HH ribozyme target se |
| | N_Geneseq_36:T49633 | + | 9.00 | 95.35 | 15 | Human CERP HH ribozyme target se |
| | N_Geneseq_36:T49635 | + | 9.00 | 95.35 | 15 | Human CERP HH ribozyme target se |
| | N_Geneseq_36:T51807 | + | 9.00 | 95.35 | 15 | Human ICAM hammerhead ribozyme t |
| | N_Geneseq_36:T50179 | + | 9.00 | 95.35 | 15 | Rabbit CERP HH ribozyme target s |
| | N_Geneseq_36:T50181 | + | 9.00 | 95.35 | 15 | Rabbit CERP HH ribozyme target s |
| | N_Geneseq_36:T50183 | + | 9.00 | 95.35 | 15 | Rabbit CERP HH ribozyme target s |
| | N_Geneseq_36:T52002 | + | 9.00 | 95.35 | 15 | Human ICAM hammerhead ribozyme t |
| | N_Geneseq_36:T52385 | + | 9.00 | 95.35 | 15 | Mouse ICAM hammerhead ribozyme t |
| | N_Geneseq_36:T52555 | + | 9.00 | 95.35 | 15 | Mouse ICAM hammerhead ribozyme t |
| | N_Geneseq_36:T72273 | + | 9.00 | 95.35 | 15 | Breast cancer mammary tumour vir |
| | N_Geneseq_36:T85295 | + | 9.00 | 95.35 | 15 | Probe 85B-DI for detection of My |
| | N_Geneseq_36:T16667 | - | 9.00 | 95.35 | 15 | Probe F67DR/0 used to identify H |
| | N_Geneseq_36:T54269 | - | 9.00 | 95.35 | 15 | Primer KC158 used in the method |
| | N_Geneseq_36:Q26132 | - | 9.00 | 94.83 | 16 | HLA-DR beta sub-type tailed prob |
| | N_Geneseq_36:Q83274 | + | 9.00 | 94.83 | 16 | c-jun antisense oligonucleotide |
| | N_Geneseq_36:Q95870 | + | 9.00 | 94.83 | 16 | Primer B (Group 11, set C) for m |
| | N_Geneseq_36:T41815 | + | 9.00 | 94.83 | 16 | HLA allele, HLA-DRB1*08, *12 and |
| | N_Geneseq_36:T90598 | + | 9.00 | 94.83 | 16 | Hepatitis C virus recognition se |
| | N_Geneseq_36:T43484 | + | 9.00 | 94.83 | 16 | HIV-1 co-receptor fusin target s |
| | N_Geneseq_36:Q26310 | + | 9.00 | 94.35 | 17 | Pro-DR probe T1 (7d = 56). Effi |
| | N_Geneseq_36:Q26112 | + | 9.00 | 94.35 | 17 | HLA-DR beta sub-type tailed prob |
| | N_Geneseq_36:Q26121 | - | 9.00 | 94.35 | 17 | HLA-DR beta sub-type tailed prob |
| | N_Geneseq_36:Q26122 | - | 9.00 | 94.35 | 17 | HLA-DR beta sub-type tailed prob |
| | N_Geneseq_36:Q26199 | - | 9.00 | 94.35 | 17 | HLA-DR beta sub-type tailed prob |
| | N_Geneseq_36:Q26233 | - | 9.00 | 94.35 | 17 | HLA-DR beta sub-type tailed prob |
| | N_Geneseq_36:Q26241 | + | 9.00 | 94.35 | 17 | HLA-DR beta sub-type tailed prob |
| | N_Geneseq_36:Q26331 | + | 9.00 | 94.35 | 17 | HLA-DR beta sub-type tailed prob |
| | N_Geneseq_36:Q89350 | + | 9.00 | 94.35 | 17 | Treponema pallidum 47 kDa surfac |
| | N_Geneseq_36:Q89351 | - | 9.00 | 94.35 | 17 | Treponema pallidum 47 kDa surfac |
| | N_Geneseq_36:Q75280 | - | 9.00 | 94.35 | 17 | Complement receptor-1 mutant oli |

| | | | | | | | | | | | |
|---|--|--|-------------------|---------|----|-------------------------------|------------|--|--|--|--|
| N_Geneseq_36:Q79801 | - | 9.00 | 94.35 | 1.2e+03 | 17 | HCV J1 NS3-NS4 domain C200 re | | | | | |
| N_Geneseq_36:T10201 | + | 9.00 | 94.35 | 1.2e+03 | 17 | Deamidating antibody heavy ch | | | | | |
| N_Geneseq_36:T00654 | - | 9.00 | 94.35 | 1.2e+03 | 17 | Antisense primer 68 for the h | | | | | |
| N_Geneseq_36:X27470 | - | 9.00 | 94.35 | 1.2e+03 | 17 | Oligo 877 for hCG alpha-subun | | | | | |
| seq_name: N_Geneseq_36:Q69882 | | | | | | | | | | | |
| seq_documentation_block: | | | | | | | | | | | |
| ID | Q69882 | standard; DNA; 13 BP. | | | | | | | | | |
| AC | Q69882: | | | | | | | | | | |
| DE | 07-MAR-1995 | (first entry) | | | | | | | | | |
| DE | XbaI/SacI | primer. | | | | | | | | | |
| KW | DNA protein-binding assay; | test sequence; screening sequence; | | | | | | | | | |
| KW | promoter; target; TATA box; | Herpes Simplex Virus; HSV; | | | | | | | | | |
| OS | origin of replication; UL9; | transcription factor; TFIID: ds. | | | | | | | | | |
| OS | Synthetic. | | | | | | | | | | |
| PN | W09411980-A. | | | | | | | | | | |
| FD | 07-JUL-1994. | | | | | | U12388. | | | | |
| PF | 20-DEC-1993; | | | | | | US-996783. | | | | |
| PR | 23-DEC-1992; | US-996783. | | | | | | | | | |
| PR | 17-SEP-1993; | US-123936. | | | | | | | | | |
| PA | (GENE-) GENELABS TECHNOLOGIES INC. | | | | | | | | | | |
| PI | Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM; | | | | | | | | | | |
| PI | WPI; 94-234711/28. | | | | | | | | | | |
| DR | Sequence-directed DNA-binding molecules - useful in | | | | | | | | | | |
| PT | pharmaceuticals and as molecular reagents | | | | | | | | | | |
| PT | Disclosure; Fig 29A; 587pp; English. | | | | | | | | | | |
| PS | A DNA protein-binding assay is provided, useful for screening | | | | | | | | | | |
| CC | libraries of synthetic or biological cpds. for their ability | | | | | | | | | | |
| CC | to bind DNA test sequences. The assay is versatile in that any | | | | | | | | | | |
| CC | number of test sequences can be tested by placing the test sequence | | | | | | | | | | |
| CC | adjacent to a defined protein-binding screening sequence. Binding | | | | | | | | | | |
| CC | of mols. to these test sequences changes the binding characteristics | | | | | | | | | | |
| CC | of the protein mol. to its cognate binding sequence. When such a mol. | | | | | | | | | | |
| CC | binds the test sequence, the equilibrium of the DNA:protein complexes | | | | | | | | | | |
| CC | is disturbed, generating changes in the concentration of free DNA probe. | | | | | | | | | | |
| CC | One application of this method is to eucaryotic general transcription | | | | | | | | | | |
| CC | factors (e.g. TFIID), where the target region is typically selected | | | | | | | | | | |
| CC | from DNA sequences adjacent to the binding site for the eucaryotic | | | | | | | | | | |
| CC | transcription factor. Numerous exemplary test sequences are given: | | | | | | | | | | |
| CC | the sequences in Q69251-731 and Q69850 correspond to promoter targets | | | | | | | | | | |
| CC | (typically, TATA box-contg. sites) for human genes and the sequences in | | | | | | | | | | |
| CC | Q69732-849 correspond to promoter targets for viral genes. The test | | | | | | | | | | |
| CC | sequences may also be randomly generated. DNA:protein interaction may | | | | | | | | | | |
| CC | be used for screening purposes, e.g. the Herpes Simplex Virus (HSV) | | | | | | | | | | |
| CC | origin of replication and UL9 (see Q69851-52, Q69865 and Q69891). | | | | | | | | | | |
| CC | The sequences given in Q69881 and Q69883-85 illustrate sample test | | | | | | | | | | |
| CC | oligonucleotides for use in the polymerase chain reaction based | | | | | | | | | | |
| CC | selection technique. | | | | | | | | | | |
| SQ | Sequence 13 BP; 3 A; 4 C; 3 G; 3 T; | | | | | | | | | | |
| alignment_scores: | | | | | | | | | | | |
| | Quality: | 9.00 | Length: | 4 | | | | | | | |
| | Ratio: | 2.250 | Gaps: | 0 | | | | | | | |
| | Percent Similarity: | 100.000 | Percent Identity: | 100.000 | | | | | | | |
| alignment_block: | | | | | | | | | | | |
| US-08-653-294-2 x Q69882 | | | | | | | | | | | |
| Align seg 1/1 to: Q69882 from: 1 to: 13 | | | | | | | | | | | |
| 7 Leu*****Arg 10 | | | | | | | | | | | |
| | | | | | | | | | | | |
| 2 CTGAGCTCTAGA 13 | | | | | | | | | | | |
| seq_name: N_Geneseq_36:T49563 | | | | | | | | | | | |
| seq_documentation_block: | | | | | | | | | | | |
| ID | T49563 | standard; DNA; 13 BP. | | | | | | | | | |
| AC | T49563: | | | | | | | | | | |
| DR | 17-MAR-1997 | (first entry) | | | | | | | | | |
| DE | XbaI/ScaI | primer to generate UL9 binding site screening sequence | | | | | | | | | |

KW Duplex DNA; target region; binding characteristic; DNA binding protein;
 KW TFIIID; transcription factor; binding site; inhibition; enhance;
 KW cancer; inherited genetic disorder; primer; amplify; ss.
 OS Synthetic.
 PN US557844-A.
 PD 26-NOV-1996.
 PF 27-JUN-1991; 723618.
 PR 27-JUN-1991; US-723618.
 PR 23-DEC-1992; US-996783.
 PR 17-SEP-1993; US-123936.
 PR 20-DEC-1993; US-171389.
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
 DR WPI; 97-020402/02.
 PT Altering binding characteristics of DNA binding proteins to duplex
 PT DNA - by attaching specific small cpd. to target region close to the
 PT protein's binding site, useful in treatment of viral disease, cancer
 PT etc
 PS Disclosure; Fig 29A; 264pp; English.
 CC The primer sequences given in T49562-63 were used to generate the
 CC generic screening sequences given in T49564-66. The general formulae
 CC represent oligonucleotides containing the UL9 recognition sequence
 CC and a variable length test site, which were used in the method
 CC of the invention. The method of the invention comprises altering
 CC the binding characteristics of a DNA-binding protein to duplex DNA.
 CC The method comprises contacting the duplex DNA with a small molecule
 CC which binds sequence-specifically to a target region, where, when the
 CC small molecule is bound to the target region, it is adjacent to, but not
 CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
 CC The small molecule is added at a concentration effective to alter the
 CC binding of the DNA binding protein, pref. TFIIID or UL9, to its binding
 CC site on the duplex DNA. The binding of the small molecule may inhibit
 CC or enhance the binding of the DNA-binding protein to its binding site
 CC (see also T63713-4312).
 SQ Sequence 13 BP; 3 A; 4 C; 3 G; 3 T;

alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-08-653-294-2 x T49563 ..
 Align seg 1/1 to: T49563 from: 1 to: 13

7 Leu*****Arg 10
 2 CTGAGCTCTAGA 13

seq_name: N_Geneseq_36:V10946
 seq_documentation_block:
 ID V10946 standard; RNA; 13 BP.
 AC V10946;
 DT 14-JUL-1998 (first entry)
 DE Human ribozyme target sequence from HLA-A exon 3 #20.
 KW Ribozyme; target; human lymphocyte antigen; HLA-A; MHC allele;
 KW major histocompatibility complex; cleavage; suppression; transplant;
 KW incompatibility; autoimmune disease; juvenile diabetes;
 KW rheumatoid arthritis; ss.
 OS Homo sapiens.
 PN WO9704087-A1.
 PD 06-FEB-1997.
 PF 18-JUL-1996; E03173.
 PR 18-JUL-1995; EP-111256.
 PA (KRUPP/) KRUPP G.
 PA (MARG/) MARGET M.
 PA (MULL/) MULLER-RUCHHOLTZ W.
 PA (WEST/) WESTPHAL E.
 PI Krupp G, Marget M, Muller-Ruchholtz W, Westphal E;
 DR WPI; 97-132628/12.

PT Ribozyme that cleaves specific MHC allele(s) - used to inhibit graft
 PT versus host reactions, to overcome blood incompatibility and to
 PS treat auto-immune disease
 PS Claim 5; Fig 1; 76pp; German.
 CC V10915-V11123 are target sequences for a novel ribozyme which cleaves
 CC specific alleles from the major histocompatibility complex (MHC). This
 CC ribozyme contains a catalytic region and a hybridisation region which is
 CC complementary to all mRNA transcribed from vertebrate genes of a specific
 CC family of closely related MHC alleles or to mRNA from a single MHC
 CC allele, and is able to cleave such mRNA. The mRNA has a target region
 CC which in case is essentially conserved in all genes of the family but
 CC differs from genes of all other MHC alleles to such a degree that no
 CC cleavage of mRNA transcribed from these other alleles occurs. This allows
 CC the selective reduction or inhibition of expression of all genes of a
 CC family or of a single gene. This ribozyme can be used for permanent or
 CC transient suppression of expression of MHC alleles, in vivo or in vitro.
 CC Specific applications are to prevent guest vs. host or host vs. guest
 CC reactions, to prevent blood incompatibilities (partic. of the ABO, rhesus
 CC and Kell systems) and to treat autoimmune diseases such as juvenile
 CC diabetes and rheumatoid arthritis. The use of this ribozyme avoids the
 CC need for immunosuppressants in transplant patients. It provides very
 CC specific reduction of particular HLA molecules that cause incompatibility
 CC between donor and recipient.
 SQ Sequence 13 BP; 4 A; 1 C; 6 G; 2 U;

alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-08-653-294-2 x V10946/rev ..
 Align seg 1/1 to reverse of: V10946 from: 1 to: 13

7 Leu*****Arg 10
 13 CTCTCAACTCGT 2

seq_name: N_Geneseq_36:Q83327
 seq_documentation_block:
 ID Q83327 standard; DNA; 14 BP.
 AC Q83327;
 DT 20-SEP-1995 (first entry)
 DE JUB-B antisense oligonucleotide.
 KW c-jun; c-fos; jun-B; neuronal injury; cell death; neoplasm;
 KW antisense; phosphorothioate; ss.
 OS Synthetic.
 PN WO9502051-A.
 PD 19-JAN-1995.
 PF 06-JUL-1994; E02218.
 PR 10-JUL-1993; EP-111059.
 PA (BIOG-) BIOGNOSTIK GES. BIOMOLEKULARE DIAGNOSTIK.
 PI Brysch W, Schlingensiepen G, Schlingensiepen K, Schlingensiepen R;
 DR WPI; 95-066896/09.
 PT Use of antisense c-jun, c-fos or jun-B nucleic acids - for
 PT preventing and treating neuronal injury, degeneration, cell death
 PT and/or neoplasms
 PS Claim 2; Page 37; 86pp; English.
 CC Antisense nucleic acid hybridising with an area of the mRNA and/or
 CC DNA comprising the genes c-jun, jun-B or c-fos, expression of which
 CC plays a causal role in neuronal injury, degeneration, cell death and/
 CC or neoplasms, can be used to prevent and treat such conditions.
 CC c-jun antisense sequences are described in Q83267-321 and Q83440-43;
 CC jun-B antisense sequences are described in Q83322-63 and Q83444-45;
 CC and c-fos antisense sequences are described in Q83364-439 and Q83446-
 CC 51. Preferably the antisense sequences are phosphorothioate
 CC oligonucleotides since these are not destroyed as fast by endogenous
 CC factors as naturally occurring molecules.
 SQ Sequence 14 BP; 4 A; 1 C; 6 G; 3 T;

alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x Q83327/rev ..

Align seg 1/1 to reverse of: Q83327 from: 1 to: 14

7 Leu*****Arg 10
 |||||
 12 CTCCTACACGA 1

seq_name: N_Geneseq_36:Q83275

seq_documentation_block:

ID Q83275 standard; DNA; 14 BP.
 AC Q83275;
 DT 19-SEP-1995 (first entry)
 DE c-jun antisense oligonucleotide.
 KW c-jun; c-fos; jun-B; neuronal injury; cell death; neoplasm;
 KW antisense; phosphorothioate; ss.
 OS Synthetic.
 PN WO9502051-A.
 PD 19-JAN-1995.
 PF 06-JUL-1994; E02218.
 PR 10-JUL-1993; E8-111059.
 PA (BIOG-) BIOGNOSTIK GES BIOMOLEKULARE DIAGNOSTIK.
 PI Brysch W, Schlingensiepen G, Schlingensiepen K, Schlingensiepen R;
 DR WPI: 95-066896/09.
 PT Use of antisense c-jun, c-fos or jun-B nucleic acids - for
 PT preventing and treating neuronal injury, degeneration, cell death
 PT and/or neoplasms
 PS Claim 2; Page 23; 86pp; English.
 CC Antisense nucleic acid hybridising with an area of the mRNA and/or
 CC DNA comprising the genes c-jun, jun-B or c-fos, expression of which
 CC plays a causal role in neuronal injury, degeneration, cell death and/
 CC or neoplasms, can be used to prevent and treat such conditions.
 CC c-jun antisense sequences are described in Q83267-321 and Q83440-43;
 CC jun-B antisense sequences are described in Q83322-63 and Q83444-45;
 CC and c-fos antisense sequences are described in Q83364-439 and Q83446-
 CC 51. Preferably the antisense sequences are phosphorothioate
 CC oligonucleotides since these are not destroyed as fast by endogenous
 CC factors as naturally occurring molecules.
 SQ Sequence 14 BP; 3 A; 2 C; 5 G; 4 T;

alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x Q83275/rev ..

Align seg 1/1 to reverse of: Q83275 from: 1 to: 14

7 Leu*****Arg 10
 |||||
 13 CTAACCTCACGT 2

seq_name: N_Geneseq_36:T76848

seq_documentation_block:

ID T76848 standard; DNA; 14 BP.
 AC T76848;
 DT 08-DEC-1997 (first entry)
 DE Probe for HGBV NTR detection.
 KW primer; probe; PCR; polymerase chain reaction; Hepatitis GB virus; HGBV;
 KW isolation; sequencing; diagnosis; treatment; acute; chronic; hepatitis;
 KW blood; organ donation; ss.

OS Synthetic.
 FH Key Location/Qualifiers
 FT modified_base 14
 FT /*tag= a
 FT /mod_base= G-carbazole-3'
 PN WO9707246-A1.
 PD 27-FEB-1997.
 PR 14-AUG-1996; U13199.
 PR 19-APR-1996; US-635309.
 PR 14-AUG-1995; US-002255.
 PA (ABBO) ABBOTT LAB.
 PI Desai SM, Joo C, Leary TP, Marshall RL, Muerhoff AS;
 PI Mushahwar IK, Simons JN;
 DR WPI: 97-165323/15.
 PT Probes and primers for detection of hepatitis GB virus - specific
 PT for 5' non-translated region and non-structural region 3
 PS Claim 2; Page 44; 56pp; English.
 CC T76826-56 are primers and probes for detection of Hepatitis GB virus.
 CC Isolation and sequencing of the HGBV genome using the primers and
 CC probes will allow design of additional probes and polypeptides which
 CC will be useful in the diagnosis and/or treatment of HGBV, both as a
 CC prophylactic and therapeutic agent. In particular, the probes could
 CC greatly enhance the ability to more accurately diagnose acute and/or
 CC chronic viral hepatitis and could provide a safer blood and organ supply
 CC by detecting non-A, non-B and non-C hepatitis in these blood and organ
 CC donations.
 SQ Sequence 14 BP; 4 A; 2 C; 6 G; 2 T;

alignment_scores:

Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x T76848/rev ..

Align seg 1/1 to reverse of: T76848 from: 1 to: 14

7 Leu*****Arg 10
 |||||
 13 TTGTCAACTCGC 2

seq_name: N_Geneseq_36:V48775

seq_documentation_block:

ID V48775 standard; DNA; 14 BP.
 AC V48775;
 DT 15-OCT-1998 (first entry)
 DE Erbb-2 gene antisense oligonucleotide Erbb-2-67.
 KW Erbb-2; antisense oligonucleotide; modulate; gene expression; ss.
 OS Synthetic.
 OS Homo sapiens.
 PN EP-856579-A1.
 PD 05-AUG-1998.
 PF 31-JAN-1997; EP-101531.
 PR 31-JAN-1997; EP-101531.
 PA (BIOG-) BIOGNOSTIK GES BIOMOLEKULARE DIAGNOSTIK.
 PI Brysch W, Schlingensiepen K;
 DR WPI: 98-400910/35.
 PT Preparation of antisense oligo:nucleotide(s) which lack long runs of
 PT consecutive guanosine or inosine - and have specific ratio of
 PT residues able to form two or three hydrogen bonds, have greater
 PT activity and reduced toxicity, used therapeutically or to modulate
 PT growth of cells in culture
 PS Claim 10; Fig 6b; 286pp; English.
 CC V48709-886 represent antisense oligonucleotides directed against the
 CC Erbb-2 gene. Of these, only oligonucleotides V48709-91 resulted
 CC in significant reduction in Erbb-2 protein expression, while
 CC oligonucleotides V48792-886 had little effect. The oligonucleotides
 CC exemplify the invention. The specification describes oligonucleotides
 CC that contain 8-30 nucleotides, which contain at most 8 nucleotides that
 CC can each form three hydrogen bonds to cytosine; do not contain four

CC consecutive nucleotides able to form three H-bonds each to four
 CC consecutive cytosines; do not contain two sequences of three consecutive
 CC nucleotides each able to form three H-bonds to three consecutive
 CC cytosines, and the ratio between residues able to form two H-bonds each
 CC (2R) or three such bonds (3R) is given by $2R/3R = 0.33-0.72$. The
 CC oligonucleotides are used to modulate expression of genes, particularly
 CC the genes for p53, ErbB-2, JunB, JunD, TGF-beta 1 or beta 2 to control
 CC proliferation of primary cell cultures (e.g. bone marrow stem, liver or
 CC kidney cells, osteoclasts, osteoblasts and/or keratinocytes). The
 CC oligonucleotides can also be used to analyse function of proteins (by
 CC altering their expression or activity) and therapeutically, e.g. in cases
 CC of cancer or (targeting TGF) for stimulating the immune system.
 SQ Sequence 14 BP; 2 A; 5 C; 3 G; 4 T;

alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x V48775 ..

Align seg 1/1 to: V48775 from: 1 to: 14

7 Leu*****Arg 10

|||||||

3 TTGTCATCCAGG 14

seq_name: N_Geneseq_36:V48580

seq_documentation_block:

ID V48580 standard; DNA; 14 BP.

AC V48580;

DE 15-OCT-1998 (first entry)

DE JunB gene antisense oligonucleotide JunB-16.

KW JunB; JunD; antisense oligonucleotide; modulate; gene expression; ss.

OS Synthetic.

OS Homo sapiens.

PN EP-856579-A1.

PD 05-AUG-1998.

PF 31-JAN-1997; 101531.

PR 31-JAN-1997; EP-101531.

PA (BIOG-) BIOGNOSTIK GES BIOMOLEKULARE DIAGNOSTIK.

PI Brysch W, Schlingensiefen K;

DR WPI: 98-400910/35

PT Preparation of antisense oligo:nucleotide(s) which lack long runs of

PT consecutive guanosine or inosine - and have specific ratio of

PT residues able to form two or three hydrogen bonds, have greater

PT activity and reduced toxicity, used therapeutically or to modulate

PT growth of cells in culture

PS Claim 10; Fig 5a; 286pp; English.

CC V48564-708 represent antisense oligonucleotides directed against the

CC JunB and JunD genes. Of these, only oligonucleotides V48565-614 resulted

CC in effective downregulation of negative growth control by JunB or

CC JunD, while V48615-708 had little effect. The oligonucleotides

CC exemplify the invention. The specification describes oligonucleotides

CC that contain 8-30 nucleotides, which contain at most 8 nucleotides that

CC can each form three hydrogen bonds to cytosine; do not contain four

CC consecutive nucleotides able to form three H-bonds each to four

CC consecutive cytosines; do not contain two sequences of three consecutive

CC nucleotides each able to form three H-bonds to three consecutive

CC cytosines, and the ratio between residues able to form two H-bonds each

CC (2R) or three such bonds (3R) is given by $2R/3R = 0.33-0.72$. The

CC oligonucleotides are used to modulate expression of genes, particularly

CC the genes for p53, ErbB-2, JunB, JunD, TGF-beta 1 or beta 2 to control

CC proliferation of primary cell cultures (e.g. bone marrow stem, liver or

CC kidney cells, osteoclasts, osteoblasts and/or keratinocytes). The

CC oligonucleotides can also be used to analyse function of proteins (by

CC altering their expression or activity) and therapeutically, e.g. in cases

CC of cancer or (targeting TGF) for stimulating the immune system.

CC Sequence 14 BP; 1 A; 6 C; 3 G; 4 T;

SQ

alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x V48580 ..

Align seg 1/1 to: V48580 from: 1 to: 14

7 Leu*****Arg 10

|||||||

3 TTGTCATCCAGG 14

seq_name: N_Geneseq_36:Q35256

seq_documentation_block:

ID Q35256 standard; DNA; 15 BP.

AC Q35256;

DE 24-MAY-1993 (first entry)

DE Plant transforming principle detection PCR primer.

KW Polymerase chain reaction; TR-DNA ROLB gene;

KW Agrobacterium rhizogenes; ss.

OS Synthetic.

PN J04356189-A.

PD 09-DEC-1992.

PF 31-MAY-1991; 128924.

PR 31-MAY-1991; JP-128924.

PA (SHMA) SHIMADZU CORP.

DR WPI: 93-Q30366/04.

PT Oligo:nucleotide for detecting plant transforming principle - is

PT gene coded to DNA of Agrobacterium rhizogenes

PS Claim 1; Page 2; 10pp; Japanese

CC The sequence is that of a PCR primer which is used as part of a

CC method for the detection of the plant transforming principle of

CC Agrobacterium rhizogenes. It detects the latter half of the TR-DNA

CC ROLB gene. The method provides highly sensitive, easy and highly

CC selective detection of a specific foreign gene from the

CC transforming principle.

SQ Sequence 15 BP; 3 A; 4 C; 6 G; 2 T;

alignment_scores:

Quality: 9.00 Length: 4

Ratio: 2.250 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x Q35256/rev ..

Align seg 1/1 to reverse of: Q35256 from: 1 to: 15

7 Leu*****Arg 10

|||||||

14 TTATCCGCCAGG 3

seq_name: N_Geneseq_36:T41816

seq_documentation_block:

ID T41816 standard; DNA; 15 BP.

AC T41816;

DE 18-DEC-1996 (first entry)

DE HLA allele, HLA-DRB1*08, *12 and *1404 resolution probe, F67.

KW Human leukocyte antigen; HLA; allele; HLA-DR*08; HLA-DR*12; locus B1;

KW polymorphism; amplification; conserved region; detection; probe;

KW tissue matching; identifying disease susceptibility; ss.

OS Synthetic.

PN U85545526-A.

PD 13-AUG-1996.

PF 27-JUN-1990; 544218.

PR 27-JUN-1990; US-544218.

PR 01-MAR-1993; US-025038.

PA (BLOO-) BLOOD CENT RES FOUND INC.
 PI Baxter-Lowe LA;
 DR Human leukocyte antigen typing of tissue samples - using
 PT allele-specific amplification to distinguish allele pairs
 PT Example 1; Column 19; 24pp; English.
 PS The sequences given in T4181-20 represent probes which were used to
 CC resolve the human leukocyte antigen (HLA) DRB1 alleles, DRB1*08, *12
 CC and *1404. This probe sequence hybridises to the Phe67 coding region
 CC found in alleles *0801, *0802, *0804, *0805 and *1202. These probes may
 CC be used in the method of invention which concerns HLA typing of a sample
 CC for an unknown pair of alleles. The pair of alleles comprises one of
 CC two known types which have the same overall set of polymorphisms but have
 CC a different distribution of polymorphisms between their two alleles. The
 CC method comprises selectively amplifying the DNA of just one allele of
 CC the unknown pair and analysing the amplified DNA to determine which
 CC polymorphisms are present in that allele, and therefore assigning the
 CC unknown pair to the known type having that allele. The method comprises
 CC three test stages. The first stage is to establish the number of alleles
 CC present in each sample. Primers corresponding to fairly well conserved
 CC regions of a locus will increase the likelihood that unknown alleles
 CC will be amplified and potentially detected by hybridisation with a
 CC probe. In the second stage, the group or basic type identified
 CC determines which set of allele specific primers will be used. The first
 CC of the two primers comprises an opt. labeled sequence common to each
 CC allele of the group identified in the first stage but different from
 CC other groups identified in stage one. The second primer may be a
 CC mixture of different labeled primers, complementary to two or more
 CC sequences within the group, or the amplification may be performed with
 CC only one second primer to detect the presence of a single group of
 CC alleles. In the third stage the specific allele is determined. This
 CC may be done by amplification or hybridisation using a radiolabelled
 CC probe. The method may be used for tissue matching, identifying disease
 CC susceptibility, etc. The method of the invention esp. distinguishes
 CC between DOB1*0304/DOB1*03032 and DOB1*0301/DOB1*0302.
 SQ Sequence 15 BP; 3 A; 4 C; 4 G; 4 T;

alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x T41816 ..

Align seg 1/1 to: T41816 from: 1 to: 15

7 Leu*****Arg 10
 |||||
 1 CTGCTCTCCAGG 12

seq_name: N_Geneseq_36:T49643

seq_documentation_block:

ID T49643 standard; RNA; 15 BP.
 AC T49643:
 DT 28-FEB-1997 (first entry)
 DE Human CETP HH ribozyme target sequence #550.
 KW Hammerhead ribozyme; cholesterol ester transfer protein; mRNA cleavage;
 KW neutral lipid transfer; plasma lipoprotein; atherosclerosis; atherectomy;
 KW reverse cholesterol transport; high density lipoprotein; therapy; CETP;
 KW familial hypercholesterolaemia; dyslipidaemia; hypoalphalipoproteinaemia;
 KW peripheral vascular disease; hyperbetalipoproteinaemia; RCT; inhibitor;
 KW angioplastic restenosis; low density lipoprotein; diabetes; HDL; human;
 KW LDL; ss.
 OS Homo sapiens.
 PN WO9620279-A1.
 PD 04-JUL-1996.
 PF 11-DEC-1995; U16000.
 PR 23-DEC-1994; US-363240.
 PA (RIBO-) RIBOZYME PHARM INC.
 PA (WARN) WARNER LAMBERT CO.

PI Bisgaier C, Couture L, McSwiggen J, Pape M, Stinchcomb D;
 DR WPI; 96-321852/32.
 PT New ribozyme(s) for cleaving cholesterol ester transfer protein mRNA
 PT - useful for preventing or treating initial development, progression
 PT or regression of vascular diseases, esp. familial
 PT hypercholesterolaemia
 PS Claim 4; Page 29; 72pp; English.
 CC T49608-T49863 represent target sequences for the human cholesterol ester
 CC transfer protein (CETP) hammerhead (HH) ribozymes (see T49881-T50137).
 CC CETP is a 74 kD glycoprotein that facilitates neutral lipid transfer
 CC between plasma lipoproteins. The numbering of the targets refers to the
 CC position of the cleavage site in full length CETP. The ribozyme binds to
 CC 5 nucleotides either side of this site, provided the sequence 5H is
 CC immediately upstream. The ribozymes are able to cleave mRNA from the
 CC gene encoding CETP, thereby blocking synthesis and/or expression of the
 CC mRNA. By inhibiting CETP, the reverse cholesterol transport (RCT)
 CC pathway can be inhibited (or eliminated) thereby preventing the reduction
 CC in size density of the high density lipoproteins (HDL), prolonging HDL
 CC half life, and therefore increasing HDL levels. The ribozymes can be
 CC used to treat conditions associated with abnormal levels of CETP,
 CC specifically familial hypercholesterolaemia, atherosclerosis, peripheral
 CC vascular disease, hyperbetalipoproteinaemia, hypoalphalipoproteinaemia,
 CC dyslipidaemia, vascular complications of diabetes, transplant
 CC atherectomy and angioplastic restenosis. By inhibiting CETP, the levels
 CC of HDL and low density lipoproteins (LDL), and the HDL:LDL ratio are
 CC favourably altered (a decrease in LDL levels, and a corresponding
 CC increase in HDL levels). The HH ribozymes can also be used
 CC diagnostically to study genetic drift and mutations in diseased cells,
 CC and to detect CETP mRNA. As the HH ribozymes target specific regions of
 CC the CETP gene, they have low non-specific activity.
 SQ Sequence 15 BP; 4 A; 5 C; 2 G; 4 U;

alignment_scores:

Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x T49643 ..

Align seg 1/1 to: T49643 from: 1 to: 15

7 Leu*****Arg 10
 |||||
 1 UUGACCUCCAGA 12

seq_name: N_Geneseq_36:T49633

seq_documentation_block:

ID T49633 standard; RNA; 15 BP.
 AC T49633:
 DT 28-FEB-1997 (first entry)
 DE Human CETP HH ribozyme target sequence #535.
 KW Hammerhead ribozyme; cholesterol ester transfer protein; mRNA cleavage;
 KW neutral lipid transfer; plasma lipoprotein; atherosclerosis; atherectomy;
 KW reverse cholesterol transport; high density lipoprotein; therapy; CETP;
 KW familial hypercholesterolaemia; dyslipidaemia; hypoalphalipoproteinaemia;
 KW peripheral vascular disease; hyperbetalipoproteinaemia; RCT; inhibitor;
 KW angioplastic restenosis; low density lipoprotein; diabetes; HDL; human;
 KW LDL; ss.
 OS Homo sapiens.
 PN WO9620279-A1.
 PD 04-JUL-1996.
 PF 11-DEC-1995; U16000.
 PR 23-DEC-1994; US-363240.
 PA (RIBO-) RIBOZYME PHARM INC.
 PA (WARN) WARNER LAMBERT CO.
 PI Bisgaier C, Couture L, McSwiggen J, Pape M, Stinchcomb D;
 DR WPI; 96-321852/32.
 PT New ribozyme(s) for cleaving cholesterol ester transfer protein mRNA
 PT - useful for preventing or treating initial development, progression
 PT or regression of vascular diseases, esp. familial

```

PT hypercholesterolaemia
PS Claim 4; Page 29; 72pp; English.
CC T49608-T49863 represent target sequences for the human cholesterol ester
CC transfer protein (CEP) hammerhead (HH) ribozymes (see T49881-T50137).
CC CEP is a 74 kD glycoprotein that facilitates neutral lipid transfer
CC between plasma lipoproteins. The numbering of the targets refers to the
CC position of the cleavage site in full length CEP. The ribozyme binds to
CC 5 nucleotides either side of this site, provided the sequence UH is
CC immediately upstream. The ribozymes are able to cleave mRNA from the
CC gene encoding CEP, thereby blocking synthesis and/or expression of the
CC mRNA. By inhibiting CEP, the reverse cholesterol transport (RCT)
CC pathway can be inhibited (or eliminated) thereby preventing the reduction
CC in size density of the high density lipoproteins (HDL), prolonging HDL
CC half life, and therefore increasing HDL levels. The ribozymes can be
CC used to treat conditions associated with abnormal levels of CEP.
CC specifically familial hypercholesterolaemia, atherosclerosis, peripheral
CC vascular disease, hyperbetalipoproteinaemia, hypopalipoproteinaemia,
CC dyslipidaemia, vascular complications of diabetes, transplant,
CC atherectomy and angioplastic restenosis. By inhibiting CEP, the levels
CC of HDL and low density lipoproteins (LDL), and the HDL:LDL ratio are
CC favourably altered (a decrease in LDL levels, and a corresponding
CC increase in HDL levels). The HH ribozymes can also be used
CC diagnostically to study genetic drift and mutations in diseased cells,
CC and to detect CEP mRNA. As the HH ribozymes target specific regions of
CC the CEP gene, they have low non-specific activity.
SQ Sequence 15 BP; 4 A; 3 C; 3 G; 5 U;

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-2 x T49633 ..
Align seg 1/1 to: T49633 from: 1 to: 15

7 Leu*****Arg 10
|||||||
2 UUGACUUCGAGA 13

seq_name: N_Geneseq_36:T49635

seq_documentation_block:
ID T49635 standard; RNA; 15 BP.
AC T49635;
DT 28-FEB-1997 (first entry)
DE Human CEP HH ribozyme target sequence #26.
KW Hammerhead ribozyme; cholesterol ester transfer protein; mRNA cleavage;
KW neutral lipid transfer; plasma lipoprotein; atherosclerosis; atherectomy;
KW reverse cholesterol transport; high density lipoprotein; therapy; CEP;
KW familial hypercholesterolaemia; dyslipidaemia; hypopalipoproteinaemia;
KW peripheral vascular disease; hyperbetalipoproteinaemia; RCT; inhibitor;
KW angioplastic restenosis; low density lipoprotein; diabetes; HDL; human;
KW LDL; ss.
OS Homo sapiens.
PN WO9620279-A1.
PD 04-JUL-1996.
PF 11-DEC-1995; U16000.
PR 23-DEC-1994; US-363240.
PA (RIBO-) RIBOZYME PHARM INC.
PA (WARN) WARNER LAMBERT CO.
PI Bisgaier C, Couture L, McSwiggen J, Pape M, Stinchcomb D;
DR WPI: 96-321852/32.
PT New ribozyme(s) for cleaving cholesterol ester transfer protein mRNA
PT - useful for preventing or treating initial development, progression
PT or regression of vascular diseases, esp. familial
PT hypercholesterolaemia
PS Claim 4; Page 29; 72pp; English.
CC T49608-T49863 represent target sequences for the human cholesterol ester
CC transfer protein (CEP) hammerhead (HH) ribozymes (see T49881-T50137).
CC CEP is a 74 kD glycoprotein that facilitates neutral lipid transfer

```

```

CC between plasma lipoproteins. The numbering of the targets refers to the
CC position of the cleavage site in full length CEP. The ribozyme binds to
CC 5 nucleotides either side of this site, provided the sequence UH is
CC immediately upstream. The ribozymes are able to cleave mRNA from the
CC gene encoding CEP, thereby blocking synthesis and/or expression of the
CC mRNA. By inhibiting CEP, the reverse cholesterol transport (RCT)
CC pathway can be inhibited (or eliminated) thereby preventing the reduction
CC in size density of the high density lipoproteins (HDL), prolonging HDL
CC half life, and therefore increasing HDL levels. The ribozymes can be
CC used to treat conditions associated with abnormal levels of CEP.
CC specifically familial hypercholesterolaemia, atherosclerosis, peripheral
CC vascular disease, hyperbetalipoproteinaemia, hypopalipoproteinaemia,
CC dyslipidaemia, vascular complications of diabetes, transplant,
CC atherectomy and angioplastic restenosis. By inhibiting CEP, the levels
CC of HDL and low density lipoproteins (LDL), and the HDL:LDL ratio are
CC favourably altered (a decrease in LDL levels, and a corresponding
CC increase in HDL levels). The HH ribozymes can also be used
CC diagnostically to study genetic drift and mutations in diseased cells,
CC and to detect CEP mRNA. As the HH ribozymes target specific regions of
CC the CEP gene, they have low non-specific activity.
SQ Sequence 15 BP; 3 A; 3 C; 4 G; 5 U;

alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-2 x T49635 ..
Align seg 1/1 to: T49635 from: 1 to: 15

7 Leu*****Arg 10
|||||||
1 UUGACUUCGAGA 12

seq_name: N_Geneseq_36:T51807

seq_documentation_block:
ID T51807 standard; RNA; 15 BP.
AC T51807;
DT 08-MAR-1997 (first entry)
DE Human ICAM hammerhead ribozyme target sequence (nt. position 31).
KW Enzymatic nucleic acid; ribozyme; trans cleavage; inhibition;
KW gene expression; downregulation; interleukin-5; IL-5; ICAM-1;
KW intercellular adhesion molecule; rel A; tumour necrosis factor;
KW TNF-alpha; respiratory syncytial virus; RSV; bcr-abl; oncogene;
KW translocation; chronic myelogenous leukaemia; CML; cancer;
KW Philadelphia chromosome; inflammation; autoimmune disease;
KW atherosclerosis; myocardial infarction; stroke; restenosis;
KW transplant rejection; rheumatoid arthritis; psoriasis;
KW myocardial ischaemia; Kawasaki disease; septic shock; HIV;
KW human immunodeficiency virus; acquired immune deficiency syndrome;
KW AIDS; ss.
OS Homo sapiens.
PN WO9523225-A2.
PD 31-AUG-1995.
PF 23-FEB-1995; IB0156.
PR 23-FEB-1994; US-201109.
PR 29-MAR-1994; US-218934.
PR 04-APR-1994; US-222795.
PR 07-APR-1994; US-224483.
PR 15-APR-1994; US-228041.
PR 15-APR-1994; US-227958.
PR 18-MAY-1994; US-245736.
PR 06-JUL-1994; US-271280.
PR 15-AUG-1994; US-291932.
PR 16-AUG-1994; US-291433.
PR 17-AUG-1994; US-292620.
PR 19-AUG-1994; US-293520.
PR 02-SEP-1994; US-300000.
PR 08-SEP-1994; US-303039.

```

```

PR 23-SEP-1994; US-311749.
PR 23-SEP-1994; US-311486.
PR 28-SEP-1994; US-314397.
PR 03-OCT-1994; US-316771.
PR 07-OCT-1994; US-319492.
PR 11-OCT-1994; US-321993.
PR 04-NOV-1994; US-334847.
PR 10-NOV-1994; US-337608.
PR 28-NOV-1994; US-345516.
PR 16-DEC-1994; US-357577.
PR 23-DEC-1994; US-363233.
PR 30-JAN-1995; US-380734.
PA (RIBO-) RIBOZYME PHARM INC.
PI Stinchcomb DT, Chowira B, Drenzo A, Draper KG, Dudycz LW;
PI Grimm S, Karpelsky A, Kisich K, Matulic-Adamic J;
PI McSwiggen JA, Modak A, Pavco P, Beigelman L, Sullivan SM;
PI Sweedler D, Thompson JD, Tracz D, Usman N, Wincott FE;
PI Woolf T;
DR WPI; 95-351090/45.
PT Ribozymes having modified bases and methods for producing them -
PT for use in inhibiting disease related genes
PS Claim 2: Page 172: 407bp: English.
CC The present sequence represents a preferred target sequence for
CC an enzymatic nucleic acid (i.e. a ribozyme) which cleaves ICAM-1
CC mRNA. Regions of the mRNA that do not form secondary folding
CC structures and that contain potential hammerhead and hairpin
CC Ribozyme cleavage sites were identified by computer analysis.
CC Ribozymes directed against these mRNA sequences were designed and
CC synthesised with modifications that improve their nuclease
CC resistance. The ribozymes cleave the ICAM-1 target sequences and
CC thereby inhibit ICAM-1 expression, making them useful for reducing
CC transplant rejection and alleviating symptoms in patients with
CC rheumatoid arthritis, asthma and other inflammatory disorders.
SQ Sequence 15 BP; 3 A; 5 C; 3 G; 4 U;

alignment_scores:
    Quality: 9.00      Length: 4
    Ratio: 2.250      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-08-653-294-2 x T51807/rev ..
Align seg 1/1 to reverse of: T51807 from: 1 to: 15

7 Leu*****Arg 10
|||||||
13 CTGAGTAGCAGA 2

seq_name: N_Geneseq_36:T50179
seq_documentation_block:
ID T50179 standard; RNA; 15 BP.
AC T50179;
DT 07-MAR-1997 (first entry)
DE Rabbit CERP HH ribozyme target sequence #372.
KW Hammerhead ribozyme; cholesterol ester transfer protein; mRNA cleavage;
KW neutral lipid transfer; plasma lipoprotein; atherosclerosis; atherectomy;
KW reverse cholesterol transport; high density lipoprotein; therapy; CERP;
KW familial hypercholesterolaemia; dyslipidaemia; hypoalipoproteinaemia;
KW peripheral vascular disease; hyperbetalipoproteinaemia; RCT; inhibitor;
KW angioplastic restenosis; low density lipoprotein; diabetes; HDL; rabbit;
KW LDL; ss.
OS Oryctolagus cuniculus.
PN WO9620279-A1.
PD 04-JUL-1996.
PF 11-DEC-1995; U16000.
PR 23-DEC-1994; US-363240.
PA (RIBO-) RIBOZYME PHARM INC.
PA (WARN) WARNER LAMBERT CO.
PI Bisgaier C, Couture L, McSwiggen J, Pape M, Stinchcomb D;
DR WPI; 96-321852/32.

PT New ribozyme(s) for cleaving cholesterol ester transfer protein mRNA
PT - useful for preventing or treating initial development, progression
PT or regression of vascular diseases, esp. familial
PT hypercholesterolaemia
PS Claim 4: Page 40: 72bp: English.
CC T50138-T50359 represent target sequences for the rabbit cholesterol ester
CC transfer protein (CERP) hammerhead (HH) ribozymes (see T50360-T50546).
CC CERP is a 74 kD glycoprotein that facilitates neutral lipid transfer
CC between plasma lipoproteins. The numbering of the targets refers to the
CC position of the cleavage site in full length CERP. The ribozyme then
CC binds to 5 nucleotides either side of this site. The ribozymes are able
CC to cleave mRNA from the gene encoding CERP, thereby blocking synthesis
CC and/or expression of the mRNA. By inhibiting CERP, the reverse
CC cholesterol transport (RCT) pathway can be inhibited (or eliminated)
CC thereby preventing the reduction in size density of the high density
CC lipoproteins (HDL), prolonging HDL half life, and therefore increasing
CC HDL levels. The ribozymes can be used to treat conditions associated
CC with abnormal levels of CERP, specifically atherosclerosis, familial
CC hypercholesterolaemia, peripheral vascular disease, dyslipidaemia,
CC hyperbetalipoproteinaemia, hypoalipoproteinaemia, vascular
CC complications of diabetes, transplant, atherectomy and angioplastic
CC restenosis. By inhibiting CERP, the levels of HDL and low density
CC lipoproteins (LDL), and the HDL:LDL ratio are favourably altered (a
CC decrease in LDL levels, and a corresponding increase in HDL levels). The
CC HH ribozymes can also be used diagnostically to study genetic drift and
CC mutations in diseased cells, and to detect CERP mRNA. As the HH
CC ribozymes target specific regions of the CERP gene, they have low
CC non-specific activity
SQ Sequence 15 BP; 4 A; 5 C; 2 G; 4 U;

alignment_scores:
    Quality: 9.00      Length: 4
    Ratio: 2.250      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-08-653-294-2 x T50179 ..
Align seg 1/1 to: T50179 from: 1 to: 15

7 Leu*****Arg 10
|||||||
1 UUGACCUCCAGA 12

```

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-2 to: EST:* out_format : pfs
 Date: Feb 8, 2000 4:02 AM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
 -DB=/cgn1_1/USPTO.spool/US08653294/runat_04022000_160700_15770/app_query.fasta.1
 -DB-EST -QMT=fastcap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.500
 -MINMATCH=0.100 -LOOCL=0.000 -LOOEXT=0.000 -CGAPOP=4.500
 -CGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosom62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -ALIGN=15 -MODE=LOCAL
 -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294
 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-2
 Query length: 10
 Database: EST:*
 Database sequences: 4538634
 Database length: 1887831382
 Search time (sec): 8553.360000

score_list:

| Sequence | Strd Orig | ZScore | Escore | Len | Documentation |
|--------------------|-----------|--------|--------|---------|--|
| gb_est21:AA921614 | + | 9.00 | 89.22 | 3.3e+04 | 25 ! AA921614 vyz1d09.r1 Stratagene m |
| gb_est15:D18733 | + | 9.00 | 88.29 | 3.3e+04 | 27 ! D18733 MUSGS01795 Mouse 3'-direct |
| gb_est20:N94347 | + | 9.00 | 88.29 | 3.8e+04 | 28 ! N94347 zb75g05.s1 Soares_senesce |
| gb_est20:AA864650 | + | 9.00 | 88.29 | 3.8e+04 | 28 ! AA864650 oh37b09.s1 NCI_CGAP_Kid |
| gb_est21:AA973948 | + | 9.00 | 88.29 | 3.8e+04 | 28 ! AA973948 oq12d02.s1 NCI_CGAP_GC4 |
| gb_est24:AI196054 | + | 9.00 | 88.29 | 3.8e+04 | 28 ! AI196054 ui68a04.y1 Sugano mouse |
| gb_est26:AI364399 | + | 9.00 | 88.29 | 3.8e+04 | 28 ! AI364399 qw37c11.x1 NCI_CGAP_Ut4 |
| gb_est34:AI1808936 | + | 9.00 | 88.29 | 3.8e+04 | 28 ! AI1808936 wh74c03.x1 NCI_CGAP_CLI |
| gb_est10:AA181661 | + | 9.00 | 87.46 | 4.2e+04 | 31 ! AA181661 zp5c05.r1 Stratagene N |
| gb_est15:AA531819 | + | 9.00 | 87.46 | 4.2e+04 | 31 ! AA479970 zv18b11.s1 Soares_NHMF |
| gb_est19:AA744598 | + | 9.00 | 87.46 | 4.2e+04 | 31 ! AA531819 TgESTz47e02.r1 TgME49 |
| gb_est20:AA847844 | + | 9.00 | 87.46 | 4.2e+04 | 31 ! AA744598 ny55h10.s1 NCI_CGAP_GCE |
| gb_est20:AA878823 | + | 9.00 | 87.46 | 4.2e+04 | 31 ! AA847844 od39a01.s1 NCI_CGAP_GCB |
| gb_est21:T67801 | + | 9.00 | 87.20 | 4.3e+04 | 32 ! AA878823 cf87c04.s1 NCI_CGAP_Li5 |
| gb_est3:R50309 | + | 9.00 | 86.71 | 4.6e+04 | 34 ! R50309 VJ61c12.s1 Soares_breast |
| gb_est7:W66110 | + | 9.00 | 86.71 | 4.6e+04 | 34 ! W66110 TgESTzy91b05.r1 TgME49 Ta |
| gb_est19:AA747497 | + | 9.00 | 86.71 | 4.6e+04 | 34 ! AA747497 nx77a12.s1 NCI_CGAP_EW1 |
| gb_est29:AI584193 | + | 9.00 | 86.71 | 4.6e+04 | 34 ! AI584193 fb2a03.x1 Zebrafish wa |
| gb_est35:AI813858 | + | 9.00 | 86.71 | 4.6e+04 | 34 ! AI813858 wk61a07.x1 NCI_CGAP_Pan |
| gb_est35:AI832311 | + | 9.00 | 86.71 | 4.6e+04 | 34 ! AI832311 at68b06.x1 Barstead col |
| gb_gss8:AQ073796 | + | 9.00 | 86.71 | 4.6e+04 | 34 ! AQ073796 EP(3)3179 Drosophila me |
| gb_est2:R10619 | + | 9.00 | 86.24 | 4.9e+04 | 36 ! R10619 YF35h11.r1 Soares_fetal1 |
| gb_est2:T64414 | + | 9.00 | 86.24 | 4.9e+04 | 36 ! T64414 Yc48e08.s1 Stratagene liv |
| gb_est4:H23949 | + | 9.00 | 86.02 | 5.0e+04 | 37 ! H23949 Yn75f10.s1 Soares_adult h |
| gb_est10:AA144860 | + | 9.00 | 86.02 | 5.0e+04 | 37 ! AA144860 mr73e06.r1 Stratagene m |
| gb_est12:AA291929 | + | 9.00 | 86.02 | 5.0e+04 | 37 ! AA291929 zt45g08.s1 Soares_ovary |
| gb_est24:AI182409 | + | 9.00 | 86.02 | 5.0e+04 | 37 ! AI182409 uc33h07.r1 Soares mouse |
| gb_est25:AI246903 | + | 9.00 | 86.02 | 5.0e+04 | 37 ! AI246903 qx31h05.x1 NCI_CGAP_Pan |
| gb_est23:AI318544 | + | 9.00 | 86.02 | 5.0e+04 | 37 ! AI318544 ta74g09.x1 NCI_CGAP_HSC |
| gb_est32:AI721631 | + | 9.00 | 86.02 | 5.0e+04 | 37 ! AI721631 fc30c09.s1 Zebrafish wa |
| gb_est2:T61852 | + | 9.00 | 85.80 | 5.2e+04 | 38 ! T61852 Yb92g06.s1 Stratagene liv |
| gb_est2:T71023 | + | 9.00 | 85.80 | 5.2e+04 | 38 ! T71023 Yc50c11.s1 Stratagene liv |
| gb_est2:T71791 | + | 9.00 | 85.80 | 5.2e+04 | 38 ! T71791 Yc64d10.s1 Stratagene liv |
| gb_est19:AA796605 | + | 9.00 | 85.80 | 5.2e+04 | 38 ! AA796605 vp31e06.r1 Barstead mou |
| gb_gss1:HSNC39C10 | + | 9.00 | 85.80 | 5.2e+04 | 39 ! HSNC39C10 Hsags DNA for trapped |
| gb_est1:D25655 | + | 9.00 | 85.59 | 5.3e+04 | 39 ! D25655 HUMGS04003 Human colon mu |
| gb_est8:C01323 | + | 9.00 | 85.38 | 5.5e+04 | 40 ! C01323 HUMGS0006349 Human adult |
| gb_est15:AA508464 | + | 9.00 | 85.38 | 5.5e+04 | 40 ! AA508464 nh66b09.s1 NCI_CGAP_Prf |
| gb_est18:AA707180 | + | 9.00 | 85.38 | 5.5e+04 | 40 ! AA707180 zj36e08.s1 Soares_fetal |
| gb_est21:AA999748 | + | 9.00 | 85.38 | 5.5e+04 | 40 ! AA999748 os54g04.s1 NCI_CGAP_Br2 |

gb_est25:AI300663 + 9.00 85.38 5.5e+04 40 ! AI300663 qo22a04.x1 NCI_CGAP_
 gb_est26:AI376705 + 9.00 85.38 5.5e+04 40 ! AI376705 tc29d03.x1 Soares_to
 gb_est37:AI974168 - 9.00 85.38 5.5e+04 40 ! AI974168 fd23a09.x1 Zebrafish

seq_name: gb_est21:AA921614

seq_documentation_block: 25 bp mRNA EST 20-APR-1998

LOCUS AA921614
 DEFINITION vyz1d09.r1 Stratagene mouse macrophage (#937306) Mus musculus CDNA
 clone IMAGE:1296113 5' similar to SW:UCP2_MOUSE P70406
 MITOCHONDRIAL UNCOUPLING PROTEIN 2 ; mRNA sequence.

ACCESSION AA921614 GI:3068393

VERSION AA921614

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 25)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE The WashU-HMHI Mouse EST Project

JOURNAL unpublished (1996)

COMMENT On May 8, 1995 this sequence version replaced gi:800982.

Contact: Marra M/Mouse EST Project

WashU-HMHI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL : contact the

IMAGE Consortium (infoimage.lnl.gov) for further information.

MGI:677161

Trace considered overall poor quality

Possible reversed clone: similarity on wrong strand

Seq primer: -28m13 rev1 ET from Amersham

High quality sequence stop: 1.

FEATURES

Location/Qualifiers

1..25

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:1296113"

/clone_lib="Stratagene mouse macrophage (#937306)"

/tissue_type="macrophage"

/dev_stage="WEHI-3 cell line"

/lab_host="SOLR (kanamycin resistant)"

/note="Organ: blood; Vector: pBluescript SK-; Site_1:

ECORI; Site_2: XhoI; Cloned unidirectionally. Primer:

Oligo dt. WEHI-3 cell line. Average insert size: 1.5 Kb;

Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG

3' -3' adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'."

BASE COUNT 4 a 5 c 10 g 6 t

ORIGIN

alignment_scores:

Quality: 9.00 Length: 4

Ratio: 2.250 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x AA921614 ..

Align seg 1/1 to: AA921614 from: 1 to: 25

7 Leu*****Arg 10

|||||||

3 TTGGCTCCAGG 14

```

seq_name: gb_est5:D18733
seq_documentation_block:
LOCUS       D18733                27 bp    mRNA                    EST          12-DEC-1995
DEFINITION  MUSGS01795 Mouse 3'-directed Mus musculus domesticus cDNA clone
            m00789 3', mRNA sequence.
ACCESSION   D18733
VERSION     D18733.1  GI:1100702
KEYWORDS    EST.
SOURCE      western European house mouse.
ORGANISM    Mus musculus domesticus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 27)
AUTHORS    Kawamoto,S., Okubo,K., Yoshii,J., Katsuki,M. and Matsubara,K.
TITLE      Analysis of gene expression in mouse embryogenesis by 3'-directed
            cDNA sequencing
JOURNAL     Unpublished (1995)
COMMENT     On May 18, 1995 this sequence version replaced gi:811454.
            Contact: Kawamoto,S., Okubo,K., Yoshii,J., Katsuki,M. and
            Matsubara,K.
            Institute for Cellular and Molecular Biology
            Osaka University
            3-1 Yamada-oka, Suita, Osaka 565, Japan
            Insert Length: 978  Std Error: 0.00
            High quality sequence stop: 234.
FEATURES             Location/Qualifiers
     source           1..27
                     /organism="Mus musculus domesticus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10092"
                     /clone="md0789"
                     /clone_lib="Mouse 3'-directed"
                     /tissue_type="decidual tissue (day 6.5-8.5 of gestation)"
BASE COUNT        10 a 5 c 6 g 6 t
ORIGIN
alignment_scores:
  Quality: 9.00      Length: 4
  Ratio: 2.250      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
US-08-653-294-2 x D18733/rev ..
Align seg 1/1 to reverse of: D18733 from: 1 to: 27
7 Leu*****Arg 10
|||||
14 CTGACTGCAAGA 3
seq_name: gb_est6:N94347
seq_documentation_block:
LOCUS       N94347                28 bp    mRNA                    EST          05-APR-1996
DEFINITION  zb75405.s1 Soares_senescent.fibroblasts_NbHSF Homo sapiens cDNA
            clone IMAGE:309464 3' similar to gb|U11120|RATRMX Rat 28S rRNA 3'
            (rRNA);, mRNA sequence.
ACCESSION   N94347
VERSION     N94347.1  GI:1266656
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 28)
AUTHORS    Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevisakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
            Wilson,R.
TITLE      The WashU-Merck EST Project

```

```

JOURNAL     Unpublished (1995)
COMMENT     On May 8, 1995 this sequence version replaced gi:799997.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL: contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: polyt not found
            Seq primer: mob.REGA+ET
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source           1..28
                     /organism="Homo sapiens"
                     /db_xref="GDB:1252878"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:309464"
                     /clone_lib="Soares_senescent.fibroblasts_NbHSF"
                     /tissue_type="senescent fibroblast"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pT73D (Pharmacia) with a modified
                     polylinker V_TYPE: phagemid; Site_1: Not 1; Site_2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dt)
                     primer 15'
                     TGTTACCAATCTGAAGTGGGCGCGCATTTTTTTTTTTT 3'
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT73 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT        6 a 7 c 7 g 7 t 1 others
ORIGIN
alignment_scores:
  Quality: 9.00      Length: 4
  Ratio: 2.250      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
alignment_block:
US-08-653-294-2 x N94347/rev ..
Align seg 1/1 to reverse of: N94347 from: 1 to: 28
7 Leu*****Arg 10
|||||
13 CTCAGTACNAGA 2
seq_name: gb_est20:AA864650
seq_documentation_block:
LOCUS       AA864650                28 bp    mRNA                    EST          13-MAY-1998
DEFINITION  ch37b09.s1 NCI_CGAP_kid6 Homo sapiens cDNA clone IMAGE:145961 3'
            similar to SW.ATPO_BOVIN P13620 ATP SYNTHASE D CHAIN, MITOCHONDRIAL
            ;, mRNA sequence.
ACCESSION   AA864650
VERSION     AA864650.1  GI:2958963
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 28)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     On Jan 17, 1998 this sequence version replaced gi:1900516.
            Contact: Robert Strausberg, Ph.D.

```

Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Stratagene, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 952 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .28
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1459961"
/clone_lib="NCI-CGAP_Kid6"
/sex="mixed"
/tissue_type="kidney tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: kidney; Vector: Bluescript SK-; Site: 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: oligo dt. Pooled kidney tumors. 5' adaptor sequence: 5' GAATTCGGCAGCAG 3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.0 kb." 4 t

BASE COUNT 8 a 5 c 11 g 4 t
ORIGIN
alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-2 x AA864650/rev ..
Align seg 1/1 to reverse of: AA864650 from: 1 to: 28

7 Leu*****Arg 10
|||||
25 CTCACCTCCAGG 14

seq_name: gb_est21:AA973948
seq_documentation_block:
LOCUS AA973948 28 bp mRNA EST 23-JUL-1998
DEFINITION oq12d02.s1 NCI-CGAP.GC4 Homo sapiens CDNA clone IMAGE:1586115 3' similar to TR:Q35787 Q35787 KINESIN-RELATED PROTEIN. ; mRNA sequence.

ACCESSION AA973948
VERSION AA973948.1 GI:3149128
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 28)

REFERENCE
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1407105.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 790 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .28

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1586115"
/clone_lib="NCI-CGAP_GC4"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from 3 pooled germ cell tumors, and was then primed with a Not I - oligo(dt) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p77T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo." 8 t

BASE COUNT 6 a 3 c 11 g 8 t
ORIGIN
alignment_scores:
Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-2 x AA973948/rev ..
Align seg 1/1 to reverse of: AA973948 from: 1 to: 28

7 Leu*****Arg 10
|||||
18 TTGACATCCAGG 7

seq_name: gb_est24:AI196054
seq_documentation_block:
LOCUS AI196054 28 bp mRNA EST 14-OCT-1998
DEFINITION ui68404.y1 Sugano mouse liver mliia Mus musculus CDNA clone IMAGE:1887534 5' similar to TR:Q99624 Q99624 TRANSPORTER PROTEIN. ; mRNA sequence.

ACCESSION AI196054
VERSION AI196054.1 GI:3748660
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 28)

REFERENCE
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellendy, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the


```

Seq primer: -40UP from Gibco
High quality sequence stop: 1.
FEATURES
  source
    1..28
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone_lib="NCI_CGAP_CLL1"
      /tissue_type="B-cell, chronic lymphocytic leukemia"
      /lab_host="DH10B"
      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dCT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCCGATGCTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT      8 a      9 c      8 g      3 t
ORIGIN
alignment_scores:
  Quality: 9.00      Length: 4
  Ratio: 2.250      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
  US-08-653-294-2 x AI808936/rev ..
  Align seg 1/1 to reverse of: AI808936 from: 1 to: 28
    7 Leu*****Arg 10
    |||||
  18 TTGAGCTGCGG 7
seq_name: gb_est10:AA181661
seq_documentation_block:
  LOCUS      AA181661      31 bp      mRNA      EST      10-MAR-1998
  DEFINITION      ZP55C05.r1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
  CDNA clone IMAGE:613352 5', similar to TR:G1006657 G1006657
  CATHEPSIN C PRECURSOR ; , mRNA sequence.
  ACCESSION      AA181661
  VERSION        AA181661.1      GI:1765144
  KEYWORDS       EST.
  SOURCE         human.
  ORGANISM       Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE      1 (bases 1 to 31)
  AUTHORS        Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
  Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
  Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
  Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE          WashU-NCI human EST project
  JOURNAL         Unpublished (1997)
  COMMENT        On May 9, 1995 this sequence version replaced gi:803081.
  Contact: Wilson RK
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  This clone is available royalty-free through LNL ; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Trace considered overall poor quality
  Possible reversed clone: similarity on wrong strand
  Insert Length: 2069 Std Error: 0.00
  Seq primer: -28M13 rev2 from Amersham
  High quality sequence stop: 1.
  Location/Qualifiers
FEATURES
  source
    1..31
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone_lib="IMAGE:613352"
      /clone_lib="Stratagene NT2 neuronal precursor 937230"
      /tissue_type="neuroepithelial cells"
      /dev_stage="Ntera-2 neuroepithelial cells"
      /lab_host="SOLR (kanamycin resistant)"
      /note="Organ: brain; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Uninduced, exponentially growing neuroepithelial
cells (Ntera-2/ci.D1). Average insert size: 1.0 Kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGCAGAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT      0 a      13 c      11 g      7 t
ORIGIN
alignment_scores:
  Quality: 9.00      Length: 4
  Ratio: 2.250      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
  US-08-653-294-2 x AA181661 ..
  Align seg 1/1 to: AA181661 from: 1 to: 31
    7 Leu*****Arg 10
    |||||
  16 CTGCTGCTGCG 27
seq_name: gb_est15:AA479970
seq_documentation_block:
  LOCUS      AA479970      31 bp      mRNA      EST      08-AUG-1997
  DEFINITION      ZV18B11.s1 Soares_NHMPu_S1 Homo sapiens CDNA clone IMAGE:753981 3'
  similar to SW:CALH_HUMAN P39060 COLLAGEN ALPHA 1(XVII) CHAIN ; ,
  mRNA sequence.
  ACCESSION      AA479970
  VERSION        AA479970.1      GI:2208121
  KEYWORDS       EST.
  SOURCE         human.
  ORGANISM       Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE      1 (bases 1 to 31)
  AUTHORS        Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
  Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
  Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
  White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE          WashU-Merck EST Project 1997
  JOURNAL         Unpublished (1997)
  COMMENT        On Sep 12, 1996 this sequence version replaced gi:1393169.
  Contact: Wilson RK
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  This clone is available royalty-free through LNL ; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Trace considered overall poor quality
  Possible reversed clone: similarity on wrong strand
  Seq primer: -41m13 fwd. ET from Amersham
  High quality sequence stop: 1.
  Location/Qualifiers
FEATURES
  source
    1..31
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone_lib="IMAGE:753981"

```

/clone_lib="Soares_NHMPu_s1"
 /tissue_type="Pooled human melanocyte, fetal heart, and
 pregnant uterus"
 /lab_host="DH10B"
 /note="Organ: mixed (see below); Vector: pT73D-Pac
 (Pharmacia) with a modified polylinker; Site_1: Not I;
 Site_2: Eco RI; Equal amounts of plasmid DNA from three
 normalized libraries (melanocyte 2NBHM, pregnant uterus
 NBHPU, and fetal heart NBH19W) were mixed, and ss circles
 were made in vitro. Following HAP purification, this DNA
 was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from pools of
 5,000 clones made from the same 3 libraries. The pools
 consisted of I.M.A.G.E. clones 260232-265223,
 340488-345479, and 484488-489479."

BASE COUNT 9 a 7 c 10 g 5 t
 ORIGIN

alignment_scores:
 Quality: 9.00 Length: 4
 Ratio: 2.250 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x AA479970/rev ..

Align seg 1/1 to reverse of: AA479970 from: 1 to: 31

7 Leu*****Arg 10

|||||

18 TTGACGCAAGG 7

seq_name: gb_est15:AA531819

seq_documentation_block:
 LOCUS AA531819 31 bp mRNA EST 22-JUL-1997
 DEFINITION TgSTz47e02.r1 TgME49 invivo Bradyzoite cDNA size selected
 Toxoplasma gondii cDNA clone tgzz47e02.r1 5' similar to TR:G603568
 G603568 HISTONE H3 ;, mRNA sequence.

ACCESSION

AA531819

VERSION

AA531819.1

KEYWORDS

EST.

SOURCE

Toxoplasma gondii.

ORGANISM

Toxoplasma gondii.

Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

Sarcocystidae; Toxoplasma.

1 (bases 1 to 31)

REFERENCE

AUTHORS

Hehl, A., Manger, J., Marra, M., Parmley, S., Sibley, L.D., Hillier, L.,

Allen, M., Bowles, L., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M.,

Le, N., Jost, S., Martin, J., Moore, B., Schellenberg, K., Steptoe, M.,

Tan, F., Theising, B., Bowers, Y., Wylie, T., Ajioke, J.A., Aslett, M.A.,

Wan, K.L., Wilson, R., Waterston, R. and Boothroyd J.C.

WashU-Stanford-PAMF-NIH Toxoplasma EST project

Unpublished (1997)

On Sep 12, 1996 this sequence version replaced gi:1397996.

Contact: Marra M

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: toxo@watson.wustl.edu

Contact Steve Parmley, PAMF (76424.16@compuserve.com) for

information on clone and library availability.

Trace considered overall poor quality

Possible reversed clone: similarity on wrong strand

High quality sequence stop: 1.

Location/Qualifiers

1. .31

/organism="Toxoplasma gondii"

/strain="ME49"

/db_xref="taxon:5811"

/clone="tgzz47e02.r1"
 /clone_lib="TgME49 invivo Bradyzoite cDNA size selected"
 /dev_stage="Bradyzoite"
 /lab_host="DH10B"

/note="Vector: Bluescript II SK-; Site_1: EcoRI; Site_2:
 NotI; Mature bradyzoites were obtained from infected mouse
 brains by percoll density centrifugation. The original
 library was constructed by Steve Parmley, Palo Alto
 Medical Foundation. cDNAs were synthesized by priming with
 oligo d(T) and directionally cloned into the EcoRI/NotI
 sites of lambda gtl1. Warning: the library contains a
 small percentage of host cDNAs derived from mouse cells.
 Inserts from this cDNA library were excised with NotI and
 EcoRI, size selected in a range of 0.7 - 2.0 kb and
 subcloned into Bluescript II SK- (Adrian Hehl, Ian Manger
 and John Boothroyd, Stanford University)"

BASE COUNT 10 a 8 c 9 g 4 t

ORIGIN

alignment_scores:

Quality: 9.00 Length: 4

Ratio: 2.250 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x AA531819/rev ..

Align seg 1/1 to reverse of: AA531819 from: 1 to: 31

7 Leu*****Arg 10

|||||

13 CTAGTCGCGC 2

seq_name: gb_est19:AA744598

seq_documentation_block:

LOCUS AA744598 31 bp mRNA EST 22-JAN-1998

DEFINITION ny25a10.s1 NC1_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1272835 3'

similar to SW:GLI4_HUMAN P10075 GLI4 PROTEIN ;, mRNA sequence.

ACCESSION AA744598

VERSION AA744598.1

KEYWORDS GI:2783362

EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 31)

REFERENCE

AUTHORS

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

On Jan 9, 1998 this sequence version replaced gi:949681.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Ph.D., Gerald Marti, M.D.

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LINL at:

www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Insert Length: 2302 Std Error: 0.00

Seq primer: -40ml3 fwd. Et from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1. .31

/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone="IMAGE:1272835"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTTTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT      3 a      13 c      10 g      5 t
ORIGIN

alignment_scores:
  Quality:      9.00      Length:      4
  Ratio:        2.250      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-08-653-294-2 x AA744598
..
Align seg 1/1 to: AA744598 from: 1 to: 31

7 Leu*****Arg 10
|||||
19 CTGCTGCACGA 30

seq_name: gb_est20:AA847844

seq documentation_block:
LOCUS      AA847844      31 bp      mRNA      EST      04-MAR-1998
DEFINITION      Od39a01.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1370280
similar to TR:092615 Q92615 MYELOBLAST KIAA0217 ;, mRNA sequence.
ACCESSION      AA847844
VERSION      AA847844.1 GI:2934362
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 31)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      On Jan 19, 1998 this sequence version replaced gi:2285157.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .31

FEATURES
source

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1370280"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTTTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT      8 a      7 c      11 g      5 t
ORIGIN

alignment_scores:
  Quality:      9.00      Length:      4
  Ratio:        2.250      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-08-653-294-2 x AA847844/rev
..
Align seg 1/1 to reverse of: AA847844 from: 1 to: 31

7 Leu*****Arg 10
|||||
15 CTGCTCCAGA 4

seq_name: gb_est20:AA878823

seq documentation_block:
LOCUS      AA878823      31 bp      mRNA      EST      25-MAR-1998
DEFINITION      O187C04.s1 NCI_CGAP_Li5 Homo sapiens cDNA clone IMAGE:1437318 3'
similar to TR:Q63627 Q63627 CTD-BINDING SR-LIKE PROTEIN RAA ;, mRNA
sequence.
ACCESSION      AA878823
VERSION      AA878823.1 GI:2987788
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 31)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      On Jan 19, 1998 this sequence version replaced gi:2152184.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
unknown library type
Trace considered overall poor quality
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .31

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1437318"
/clone_lib="NCI_CGAP_Li5"
/tissue_type="hepatic adenoma"
/lab_host="DH10B"
/notes="Organ: liver; Vector: pCMV-SPORT4; Site_1: SalI;

```

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 0.8 kb."

6 a 4 c 20 g 1 t

BASE COUNT
ORIGIN

alignment_scores:

Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x AA878823/rev ..

Align seg 1/1 to reverse of: AA878823 from: 1 to: 31

7 Leu*****Arg 10

|||||

23 CTCCTCTCCGC 12

seq_name: gb_est2:T67801

seq_documentation_block:

LOCUS T67801 32 bp mRNA EST 22-FEB-1995
DEFINITION YC39a09.s1 Stragatene liver (#937224) Homo sapiens CDNA clone
IMAGE:83032 3' similar to gb:X02162 APOLIPOPROTEIN A-I PRECURSOR
(HUMAN);, mRNA sequence.

ACCESSION T67801

VERSION T67801.1 GI:678949

KEYWORDS EST.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.,
and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

97044478

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence starts: 1

High quality sequence stops: 1

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Trace considered overall poor quality

Seq primer: -2lml3

High quality sequence stop: 1.

Location/Qualifiers

1..32

/organism="Homo sapiens"

/db_xref="GDB:500089"

/db_xref="taxon:9606"

/clone="IMAGE:83032"

/clone_lib="Stragatene liver (#937224)"

/sex="male"

/dev_stage="49.years old"

/lab_host="SOLR cells (kanamycin resistant)"

/note="Organ: liver; Vector: pBluescript SK; Site: 1;

EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:

Oligo dt. Hepatotomy from normal male caucasian. Average

insert size: 1.1 kb; Uni-ZAP XR Vector; -5' adaptor

sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3'

5 a 7 c 14 g 6 t

BASE COUNT
ORIGIN

alignment_scores:

Quality: 9.00 Length: 4
Ratio: 2.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-2 x T67801 ..

Align seg 1/1 to: T67801 from: 1 to: 32

7 Leu*****Arg 10

|||||

8 CTGAGCACCCG 19

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:29:34 ; Search time 122.56 Seconds
(without alignments)
1.933 Million cell updates/sec

Title: US-08-653-294-3
Perfect score: 25
Sequence: 1 REXLRXXXX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|--------|---------------------|
| 1 | 19 | 76.0 | 10 | R41209 | Peptide fragment o |
| 2 | 19 | 76.0 | 10 | R83061 | HLA-B*7 CTL modul |
| 3 | 19 | 76.0 | 10 | W07515 | T-cell modulating |
| 4 | 19 | 76.0 | 10 | W33786 | Peptide B7.75-84 t |
| 5 | 19 | 76.0 | 10 | W33796 | Peptide B7.75-84 t |
| 6 | 19 | 76.0 | 20 | R92913 | HLA-B*7 CTL modul |
| 7 | 19 | 76.0 | 20 | R95415 | HLA-B*7.84-75-84 Pa |
| 8 | 19 | 76.0 | 20 | W33790 | Peptide B7.84-75/7 |
| 9 | 19 | 76.0 | 20 | W33797 | Peptide B7.84-75/7 |
| 10 | 19 | 76.0 | 25 | R41207 | Peptide fragment o |
| 11 | 19 | 76.0 | 25 | R83073 | HLA-B*62 CTL modul |
| 12 | 19 | 76.0 | 25 | R95431 | HLA-B*7.60-84. Comp |
| 13 | 19 | 76.0 | 25 | R95419 | HLA-B*62.60-84. Co |
| 14 | 19 | 76.0 | 62 | W00210 | Human MORF-1 prote |
| 15 | 19 | 76.0 | 95 | W85049 | Amino acid sequenc |
| 16 | 19 | 76.0 | 95 | W67637 | MSRV-1 virus clone |
| 17 | 19 | 76.0 | 97 | Y12165 | Human 5' EST secre |
| 18 | 19 | 76.0 | 110 | Y11295 | S. pneumoniae poss |
| 19 | 19 | 76.0 | 127 | Y11165 | S. pneumoniae 50S |
| 20 | 19 | 76.0 | 137 | W38577 | S. pneumoniae 50S |
| 21 | 19 | 76.0 | 201 | W87491 | Amino acid sequenc |
| 22 | 19 | 76.0 | 201 | W87492 | Amino acid sequenc |
| 23 | 19 | 76.0 | 204 | W14700 | Nucleic acid bindi |
| 24 | 19 | 76.0 | 208 | W03653 | FADD (Fas-associat |
| 25 | 19 | 76.0 | 208 | W87493 | Amino acid sequenc |
| 26 | 19 | 76.0 | 208 | W87493 | Human FADD protein |
| 27 | 19 | 76.0 | 210 | Y07103 | Colon cancer assoc |
| 28 | 19 | 76.0 | 216 | R72605 | Human calcium chan |
| 29 | 19 | 76.0 | 216 | W63157 | Human calcium chan |
| 30 | 19 | 76.0 | 219 | R72606 | Human calcium chan |
| 31 | 19 | 76.0 | 219 | W63158 | Human calcium chan |
| 32 | 19 | 76.0 | 240 | R39697 | Myasthenic antigen |
| 33 | 19 | 76.0 | 248 | R30191 | UGT1D Exon 1 produ |
| 34 | 19 | 76.0 | 256 | R98346 | MORT-1 modulator o |

35 19 76.0 256 1 W11894 Modulator of cellu
36 19 76.0 273 1 W82006 Human adult brain
37 19 76.0 283 1 W36025 Fragment of MSRV-1
38 19 76.0 283 1 W71085 Multiple sclerosis
39 19 76.0 289 1 R30190 UGT1E Exon 1 produ
40 19 76.0 292 1 R29312 Cyclin D3 protein.
41 19 76.0 292 1 R44804 Human cyclin D3. N
42 19 76.0 334 1 R28073 Brevibacterium fla
43 19 76.0 339 1 R95602 STR3 (suppressor o
44 19 76.0 340 1 R06830 Thermopsin. Gene e
45 19 76.0 350 1 W14532 Human chimeric fuc

ALIGNMENTS

RESULT 1

R41209
ID R41209 standard; peptide; 10 AA.
AC R41209;
DT 15-MAR-1994 (first entry)
DE Peptide fragment of Class I HLA peptide.
KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;
KW parasitic disease; cytotoxic T lymphocyte; modulation.
OS Synthetic.
PN W09317699-A.
PD 16-SEP-1993.
PF 25-FEB-1993; U01758.
PR 02-MAR-1992; US-844716.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger CA, Krensky AM;
DR WPI; 93-303134/38.
PT New peptide(s) based on Class I HLA antigen domains - used for
PT modulating cytotoxic T-lymphocyte activity towards targets
PS Claim 11; Page 54; 61pp; English.
CC The peptide is used to modulate cytotoxic T-lymphocyte (CTL)
CC activity, either by inhibition or stimulation. It can be used
CC for inhibiting CTL toxicity in transplantations, for inducing CTL
CC activity in parasitic diseases and neoplasia and in studies on viral
CC infection. The peptide can also be used for identifying CTLs which
CC bind to it and removing subsets of CTLs from a T-cell composition.
CC This peptide sequence is more commonly found within larger peptide
CC compounds of not more than 30 amino acids in length.
SQ Sequence 10 AA;

Query Match 76.0%; Score 19; DB 1; Length 10;

Best Local Similarity 80.0%; Pred. No. 23;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 REXLR 5

Db 1 RESLR 5

RESULT 2

R83061
ID R83061 standard; peptide; 10 AA.

AC R83061;

DT 15-MAY-1996 (first entry)

DE HLA-B*7 CTL modulating peptide (B7.75-84).

KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;

KW immunosuppressant; graft versus host disorder; transplantation; therapy;

KW class I MHC; HLA-B*7.

OS Synthetic.

PN W09526979-A1.

PD 12-OCT-1995.

PF 05-APR-1995; U04349.

PR 05-APR-1994; US-222851.

PA (STRD) UNIV LELAND STANFORD JUNIOR.

PI Clayberger C, Krensky AM, Farham P;

DR WPI; 95-358582/46.

PT Extension of acceptance period of transplants from MHC unmatched

PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 PS Claim 13; Page 66; 80pp; English.
 CC This sequence represents a fragment of a class I major histocompatibility
 CC complex (MHC) antigen. This sequence corresponds to residues 75-84 of
 CC the alpha-1 domain of the class I MHC HLA-B*7. This sequence, and the
 CC peptide fragments represented by R83062-R83085, R83090-R83096 and
 CC R82907-R82913 can be used to extend the period of acceptance by a
 CC recipient of a transplant from an MHC unmatched donor. The peptides are
 CC administered to a patient in conjunction with a subtherapeutic amount of
 CC an immunosuppressant. This is administered to the patient for a limited
 CC period of time (compared to the lifetime administration for current
 CC treatments). The peptides particularly modulate (or inhibit) the
 CC activity of the cytotoxic T lymphocytes (CTLs) of the patient.
 SQ Sequence 10 AA;

Query Match 76.0%; Score 19; DB 1; Length 10;
 Best Local Similarity 80.0%; Pred. No. 23;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5

DB 1 RESLR 5

RESULT 3

ID W07515 standard; peptide; 10 AA.

AC W07515;

DT 04-AUG-1997 (first entry)

DE T-cell modulating peptide #4

KW mammal; major histocompatibility complex; tissue destruction; alaph-domain;
 KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;
 KW rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;
 KW thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;
 KW autologous target cell; cytokine release; T cell activation; therapy.
 OS Synthetic.

PN W09635443-A1.

PD 14-NOV-1996.

PF 05-APR-1996; U04710.

PR 12-MAY-1995; US-440504.

PA (SANG-) SANGSTAT MEDICAL CORP.

PI Buelow R;

DR WPI: 96-518410/51.

PT Treatment of auto-immune disease by admin. of peptide(s) corresp. to
 PT major histocompatibility complex antigens - esp. for delaying onset
 PT of clinical symptoms of insulin dependent diabetes by modulating T
 PT cell mediated attack on target cells

PS Claim 7; Page 20; 24pp; English.

CC W07512-W07518 represent T-cell modulating peptides that can be used in
 CC the method of the invention. These sequences are based on a portion of
 CC the generic peptide corresponding to residues 70-91 of the alpha-1 domain
 CC of the major histocompatibility complex (MHC) class I antigen (see
 CC W07510). The method is for affecting the course of an autoimmune disease
 CC involving T-cell mediated destruction of tissue in mammals. These

CC peptides are used especially to treat insulin-dependent diabetes
 CC mellitus, preferably being administered during the pre-clinical stage to
 CC delay onset of the disease. Other diseases that can be treated are
 CC multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus vulgaris,
 CC Sjogren's disease, thyroid disease, Hashimoto's thyroiditis, myasthenia
 CC gravis, etc. The peptides modulate T-cell mediated attack on autologous
 CC target cells, and may also reduce inflammation, swelling, and release of
 CC cytokines, perforins, granzymes etc. associated with T cell activation.
 SQ Sequence 10 AA;

Query Match

Best Local Similarity 76.0%; Score 19; DB 1; Length 10;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5

DB 1 RESLR 5
 || || ||

RESULT 4

ID W33786 standard; peptide; 10 AA.

AC W33786;

DT 19-JUN-1998 (first entry)

DE Peptide B7.75-84 tested for immunomodulating activity.

KW Immunomodulating dimer; immunosuppressant drug; CTL activation;

KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;

KW rejection.

OS Synthetic.

OS Homo sapiens.

PN W09744351-A1.

PD 27-NOV-1997.

PF 22-MAY-1997; U08689.

PR 24-MAY-1996; US-653294.

PA (STRD) UNIV LELAND STANFORD JUNIOR.

PI Buelow R; Clayberger C; Krensky AM;

DR WPI: 98-086530/08.

PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B

PT alpha-1 domain, used for preventing rejection of transplants or

PT treating autoimmune diseases

PS Example 1; Page 19; 41pp; English.

CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
 CC activity. A peptide-type compound or variant is claimed which has
 CC immunomodulating activity, including the N-terminal acylated and/or
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where
 CC the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha-1 domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 SQ Sequence 10 AA;

Query Match 76.0%; Score 19; DB 1; Length 10;

Best Local Similarity 80.0%; Pred. No. 23;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5

DB 1 RESLR 5

RESULT 5

ID W33796 standard; peptide; 10 AA.

AC W33796;

DT 19-JUN-1998 (first entry)

DE Peptide B7.75-84 tested for immunomodulating activity.

KW Immunomodulating dimer; immunosuppressant drug; CTL activation;

KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;

KW rejection.

OS Synthetic.

OS Homo sapiens.

PN W09744351-A1.

PD 27-NOV-1997.

PF 22-MAY-1997; U08689.

PR 24-MAY-1996; US-653294.

PA (STRD) UNIV LELAND STANFORD JUNIOR.

PI Beulow R, Clayberger C, Krensky AM;
 DR WPI: 98-086530/08;
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Example 1; Page 19; 41pp; English.
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
 CC activity. A peptide-type compound or variant is claimed which has
 CC immunomodulating activity, including the N-terminal acylated and/or
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where
 CC the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha-1 domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 SQ Sequence 10 AA;

Query Match 76.0%; Score 19; DB 1; Length 10;
 Best Local Similarity 80.0%; Pred. No. 23;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 REXLR 5
 || ||
 DB 1 RESLR 5

RESULT 6

ID R2913 standard; peptide; 20 AA.
 AC R2913;
 DE HLA-B7 CTL modulating peptide (B7.84-75/75-84).
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW Class I MHC; HLA-B7.
 OS Synthetic.
 PN W09526979-A1.
 PD 12-OCT-1995.
 PF 05-APR-1995; U04349.
 PR 05-APR-1994; US-222851.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM, Parham P;
 DR WPI: 95-358582/46.
 PT Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 PT host
 PS Example 15; Page 36; 80pp; English.
 CC R83061-R83085, R83090-R83096 and R29207-R29213 represent fragments of
 CC class I major histocompatibility complex (MHC) antigens. This sequence
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class
 CC I MHC HLA-B7. These sequences can be used to extend the period of
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.
 CC The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 SQ Sequence 20 AA;

Query Match 76.0%; Score 19; DB 1; Length 20;

Best Local Similarity 80.0%; Pred. No. 46;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 REXLR 5
 || ||
 DB 1 RESLR 15

RESULT 7

R95415 standard; peptide; 20 AA.
 ID R95415;
 AC R95415;
 DE HLA-B7.84-75-84 Palindrome.
 KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; Calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KW cytolysis; antigen presenting cell.
 OS Synthetic.
 PN W09513288-A1.
 PD 18-MAY-1995.
 PF 10-NOV-1994; U12985.
 PR 10-NOV-1993; US-150493.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI: 95-194027/25.
 PT Compsns. comprising lymphoid surface membrane proteins - which may
 PT inhibit cytolytic activity and differentiation of CTLs.
 PS Example: Page 18; 25pp; English.
 CC R95413, and R95415-R95431 represent palindromes and fragments of
 CC human-leucocyte-associated antigens. These sequences can be used to isolate
 CC HLA-B7.84-75/75-84 palindrome. These sequences can be used to isolate
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
 CC protein associated with T-cell activation in mammalian T-cells, and is
 CC also immunologically cross reactive with the heat shock protein Hsc70.
 CC p74 is found in a limited number of cell types, but is particularly
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable
 CC cell with an amphoteric detergent, and then passed through an affinity
 CC column containing a covalently bound HLA-B2702 palindromic peptide.
 CC Compositions comprising the extracellular fragment of p74 combined with
 CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate
 CC compounds can be screened for their effect on the cytolytic activity of
 CC T-cells, by combining them with the extracellular portion of p74 and
 CC determining the amount of binding between the candidate compound and p74.
 CC Modulation of CTL activity can be inhibited in a cellular composition
 CC containing T-cells and antigen presenting cells (APCs), by adding to the
 CC mix the extracellular portion of p74, in an amount sufficient to compete
 CC with p74 for the binding of the p74 ligand.
 SQ Sequence 20 AA;

Query Match 76.0%; Score 19; DB 1; Length 20;
 Best Local Similarity 80.0%; Pred. No. 46;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 REXLR 5
 || ||
 DB 1 RESLR 15

RESULT 8

W33790 standard; peptide; 20 AA.
 ID W33790;
 AC W33790;
 DE Peptide B7.84-75/75-84 tested for immunomodulating activity.
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN W09744351-A1.

PD 27-NOV-1997. U08689.
 PF 22-MAY-1997; US-653294.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 FI Beulow R, Clayberger C, Krensky AM;
 DR WPI; 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Example 1; Page 19; 41pp; English.
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
 CC activity. A peptide-type compound or variant is claimed which has
 CC immunomodulating activity, including the N-terminal acylated and/or
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where
 CC the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 CC Sequence 20 AA;
 SQ

Query Match 76.0%; Score 19; DB 1; Length 20;
 Best Local Similarity 80.0%; Pred. No. 46;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 ||||
 DB 11 RESLR 15

RESULT 9
 W33797
 ID W33797 standard; peptide; 20 AA.
 AC W33797;
 DT 19-JUN-1998 (first entry)
 DE Peptide 87.84-75/75-84 tested for immunomodulating activity.
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN W09744351-A1.
 PD 27-NOV-1997.
 PF 22-MAY-1997; U08689.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI; 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Example 1; Page 19; 41pp; English.
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
 CC activity. A peptide-type compound or variant is claimed which has
 CC immunomodulating activity, including the N-terminal acylated and/or
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where
 CC the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise

CC amino acid sequences related to a Class I HLA-B alpha domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 CC Sequence 20 AA;
 SQ

Query Match 76.0%; Score 19; DB 1; Length 20;
 Best Local Similarity 80.0%; Pred. No. 46;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 ||||
 DB 11 RESLR 15

RESULT 10
 R41207
 ID R41207 standard; peptide; 25 AA.
 AC R41207;
 DT 15-MAR-1994 (first entry)
 DE Peptide fragment of Class I HLA peptide.
 KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;
 KW parasitic disease; cytotoxic T lymphocyte; modulation.
 OS Synthetic.
 PN W09317699-A.
 PD 16-SEP-1993.
 PF 25-FEB-1993; U01758.
 PR 02-MAR-1992; US-844716.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger CA, Krensky AM;
 DR WPI; 93-303134/38.

PT New peptide(s) based on Class I HLA antigen domains - used for
 PT modulating cytotoxic T-lymphocyte activity towards targets
 PS Claim 10; Page 54; 61pp; English.
 CC The peptide (or a fragment of at least 10 amino acids, joined at at
 CC least one terminus to a sequence other than that of wild type HLA
 CC antigen) is used to modulate cytotoxic T-lymphocyte (CTL) activity,
 CC either by inhibition or stimulation. It can be used for
 CC inhibiting CTL toxicity in transplantations, for inducing CTL
 CC activity in parasitic diseases and neoplasia and in studies on viral
 CC infection. The peptide can also be used for identifying CTLs which
 CC bind to it and removing subsets of CTLs from a T-cell composition.
 SQ Sequence 25 AA;

Query Match 76.0%; Score 19; DB 1; Length 25;
 Best Local Similarity 80.0%; Pred. No. 56;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 ||||
 DB 16 RESLR 20

RESULT 11
 R83073
 ID R83073 standard; peptide; 25 AA.
 AC R83073;
 DT 16-MAY-1996 (first entry)
 DE HLA-Bw62 CTL modulating peptide (Bw62.60-84).
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW Class I MHC; HLA-Bw62.
 OS Synthetic.
 PN W09526979-A1.
 PD 12-OCT-1995.
 PF 05-APR-1995; U04349.

PR 05-APR-1994; US-222851.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM, Parham P;
 DR WPI; 95-358582/46.
 PT Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B7-84 MHC antigen of the recipient
 PT host
 PS Example 13; Page 32; 80pp; English.
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of
 CC class I major histocompatibility complex (MHC) antigens. This sequence
 CC corresponds to residues 60-84 of the alpha-1 domain of the class I MHC
 CC HLA-B*62. These sequences can be used to extend the period of acceptance
 CC by a recipient of a transplant from an MHC unmatched donor. The peptides
 CC are administered to a patient in conjunction with a subtherapeutic amount
 CC of an immunosuppressant. This is administered to the patient for a
 CC limited period of time (compared to the lifetime administration for
 CC current treatments). The peptides particularly modulate (or inhibit) the
 CC activity of the cytotoxic T lymphocytes (CTLs) of the patient.
 SQ Sequence 25 AA;

Query Match 76.0%; Score 19; DB 1; Length 25;
 Best Local Similarity 80.0%; Pred. No. 56;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 || ||
 DB 16 RESLR 20

RESULT 12
 R95431
 ID R95431 standard; peptide; 25 AA.
 AC R95431;
 DT 12-NOV-1996 (first entry)
 DE HLA-B*62.60-84.
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KW cytolysis; antigen presenting cell.
 OS Synthetic.
 PN W09513288-A1.
 PD 18-MAY-1995.
 PF 10-NOV-1994; U12985.
 PR 10-NOV-1993; US-150493.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI; 95-194027/25.
 PT Comps. comprising lymphoid surface membrane proteins - which may
 PT inhibit cytolytic activity and differentiation of CTLs.
 PS Example; Page 12; 29pp; English.
 CC R95413, and R95415-R95431 represent palindromes and fragments of
 CC human-leucocyte-associated antigens. This sequence represents the
 CC HLA-B*62.60-84. These sequences can be used to isolate the protein p74
 CC from a T-cell lysate. p74 is a T-cell surface membrane protein
 CC associated with T-cell activation in mammalian T-cells, and is also
 CC immunologically cross reactive with the heat shock protein Hsc70. p74 is
 CC found in a limited number of cell types, but is particularly expressed on
 CC B and T cells. p74 can be isolated by lysis of a suitable cell with an
 CC amphoteric detergent, and then passed through an affinity column
 CC containing a covalently bound HLA-B2702 palindromic peptide.
 CC Compositions comprising the extracellular fragment of p74 combined with
 CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate
 CC compounds can be screened for their effect on the cytolytic activity of
 CC T-cells, by combining them with the extracellular portion of p74 and
 CC determining the amount of binding between the candidate compound and p74.
 CC Modulation of CTL activity can be inhibited in a cellular composition
 CC containing T-cells and antigen presenting cells (APCs), by adding to the
 CC mix the extracellular portion of p74, in an amount sufficient to compete
 CC with p74 for the binding of the p74 ligand.
 SQ Sequence 25 AA;

Query Match 76.0%; Score 19; DB 1; Length 25;
 Best Local Similarity 80.0%; Pred. No. 56;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 || ||
 DB 16 RESLR 20

RESULT 13
 R95419
 ID R95419 standard; peptide; 25 AA.
 AC R95419;
 DT 12-NOV-1996 (first entry)
 DE HLA-B*62.60-84.
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KW cytolysis; antigen presenting cell.
 OS Synthetic.
 PN W09513288-A1.
 PD 18-MAY-1995.
 PF 10-NOV-1994; U12985.
 PR 10-NOV-1993; US-150493.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI; 95-194027/25.
 PT Comps. comprising lymphoid surface membrane proteins - which may
 PT inhibit cytolytic activity and differentiation of CTLs.
 PS Example; Page 9; 29pp; English.
 CC R95413, and R95415-R95431 represent palindromes and fragments of
 CC human-leucocyte-associated antigens. This sequence represents the
 CC HLA-B*62.60-84. These sequences can be used to isolate the protein p74
 CC from a T-cell lysate. p74 is a T-cell surface membrane protein
 CC associated with T-cell activation in mammalian T-cells, and is also
 CC immunologically cross reactive with the heat shock protein Hsc70. p74 is
 CC found in a limited number of cell types, but is particularly expressed on
 CC B and T cells. p74 can be isolated by lysis of a suitable cell with an
 CC amphoteric detergent, and then passed through an affinity column
 CC containing a covalently bound HLA-B2702 palindromic peptide.
 CC Compositions comprising the extracellular fragment of p74 combined with
 CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate
 CC compounds can be screened for their effect on the cytolytic activity of
 CC T-cells, by combining them with the extracellular portion of p74 and
 CC determining the amount of binding between the candidate compound and p74.
 CC Modulation of CTL activity can be inhibited in a cellular composition
 CC containing T-cells and antigen presenting cells (APCs), by adding to the
 CC mix the extracellular portion of p74, in an amount sufficient to compete
 CC with p74 for the binding of the p74 ligand.
 SQ Sequence 25 AA;

Query Match 76.0%; Score 19; DB 1; Length 25;
 Best Local Similarity 80.0%; Pred. No. 56;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 || ||
 DB 16 RESLR 20

RESULT 14
 W00210
 ID W00210 standard; peptide; 62 AA.
 AC W00210;
 DT 17-APR-1997 (first entry)
 DE Human MORT-1 protein death domain motif.
 KW Death domain; regulatory protein; NGF-R; nerve growth receptor;
 KW FAS-R; Fas ligand receptor; Fas/ABO1; ankyrin 1; p35 TNF-R;
 KW tumour necrosis factor receptor; MORT1; cell cytotoxicity; HIV;
 KW human immunodeficiency virus; cancer; neoplasia; disease.

OS Homo sapiens.
 PN W09625941-AL.
 PD 29-AUG-1996.
 PF 15-FEB-1996; U02326.
 PR 22-FEB-1996; IL-112742.
 PR 13-SEP-1995; IL-115289.
 PA (WEIN/) WEINURZEL H.
 PA (YEDA) YEDA RES & DEV CO LTD.
 PI Boidin MP, Goncharov TM, Mett I, Pancer Z, Varfolomeev EE;
 PI Wallach D;
 PT WPI; 96-402125/40.
 PT Modulator of regulatory cellular events mediated by "death domain"
 PT contg. regulatory proteins - useful for modulating functions
 PT mediated in cells by proteins contg the death domain
 PS Claim 9; Fig 1; 74pp; English.
 CC W00210 shows the death domain of the MORT-1, which binds to the
 CC intracellular portion of the human Fas-ligand receptor (FAS-R). The
 CC death domain (DD) of FAS-R, p55 tumour necrosis factor receptor
 CC (p55 TNF-R), nerve growth factor receptor (NGF-R) and ankyrin 1 are also
 CC given (see W00206 and W00208-W00210). These DDs are used to identify
 CC compounds capable of modulating activity of the regulatory proteins
 CC (p55, NGF, TNF and FAS-R ligand, MORT-1) via interaction with the DDs.
 CC Such modulators which may be antibodies, antisense sequences or
 CC ribozymes (which can affect the cellular mRNA sequences encoding the
 CC proteins) and are useful for modulation of effects of the regulatory
 CC proteins within the cell. Tumour cells, HIV-infected cells or other
 CC diseased cells can be treated by targeting the cells with animal viral
 CC vectors encoding the modulators and a viral surface antigen capable of
 CC binding to a specific receptor. The DDs are characterised by having
 CC groups of common amino acid residues Trp, Ala, Asp, Glu, Thr, Arg and
 CC Tyr within locations that can be aligned to show homology.
 SQ Sequence 62 AA;

Query Match 76.0%; Score 19; DB 1; Length 62;
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 ||||
 DB 31 RESLR 35

RESULT 15
 W85049
 ID W85049 standard; Protein; 95 AA.
 AC W85049;
 DT 11-FEB-1999 (first entry)
 DE Amino acid sequence of clone GM3.
 KW Multiple sclerosis; MS; antibody fixation;
 KW multiple sclerosis-associated retrovirus; MSRV-1; vaccine.
 OS Synthetic.
 OS Multiple sclerosis-associated retrovirus.
 PN FR2762601-AL.
 PD 30-OCT-1998.
 PF 31-DEC-1997; 016870.
 PR 29-APR-1997; FR-005679.
 PA (INMR) BIO MERIEUX.
 PI Jolivet RC, Mandrand B, Perron H;
 DR WPI; 98-586098/50.
 DR N-PSDB; V63612.
 PT Peptides reactive with multiple sclerosis antibodies - useful for
 PT diagnosis or vaccine preparation
 PS Example 4; Fig 4; 43pp; French.
 CC The present sequence is encoded by clone GM3 and represents a region
 CC from a protein of multiple sclerosis-associated retrovirus MSRV-1.
 CC Peptides that specifically react with antibodies from multiple sclerosis
 CC (MS) patients can be derived from the sequence. The peptides can be used
 CC to diagnose or monitor multiple sclerosis, and to detect the
 CC multiple sclerosis-associated retrovirus MSRV-1. The peptides can
 CC also be used for "fixation" of multiple sclerosis or MSRV-1 antibodies
 CC in biological samples, especially serum, cerebrospinal fluid or urine.
 CC Vaccines can also be prepared from the peptides.

SQ Sequence 95 AA;
 Query Match 76.0%; Score 19; DB 1; Length 95;
 Best Local Similarity 80.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 REXLR 5
 ||||
 DB 16 REALR 20

Search completed: February 8, 2000, 01:29:35
 Job time: 1747 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 11:54:11 ; Search time 117.7 seconds
(without alignments)
4.008 Million cell updates/sec

Title: US-08-653-294-3
Perfect score: 25
Sequence: 1 REXLRXXXXX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : PIR62:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 19 | 76.0 | 35 | 2 D69330 | hypothetical prote |
| 2 | 19 | 76.0 | 60 | 2 S01800 | pepsin (EC 3.4.23. |
| 3 | 19 | 76.0 | 65 | 2 E70644 | probable ribosomal |
| 4 | 19 | 76.0 | 100 | 2 S44892 | 2Kil2.4 protein - |
| 5 | 19 | 76.0 | 102 | 2 S19733 | hypothetical prote |
| 6 | 19 | 76.0 | 104 | 2 C75046 | hypothetical prote |
| 7 | 19 | 76.0 | 106 | 2 S26322 | Ig heavy chain v r |
| 8 | 19 | 76.0 | 106 | 2 B44915 | hoxl protein - Azo |
| 9 | 19 | 76.0 | 106 | 2 S74019 | hypothetical prote |
| 10 | 19 | 76.0 | 124 | 2 C24733 | myosin alpha heavy |
| 11 | 19 | 76.0 | 130 | 2 C72247 | ribosomal protein |
| 12 | 19 | 76.0 | 130 | 2 JQ0370 | photosystem I ilk |
| 13 | 19 | 76.0 | 135 | 2 S78260 | ribosomal protein |
| 14 | 19 | 76.0 | 137 | 1 R5M16 | ribosomal protein |
| 15 | 19 | 76.0 | 137 | 2 I38875 | MHC class I antige |
| 16 | 19 | 76.0 | 137 | 2 I80176 | class I histocompa |
| 17 | 19 | 76.0 | 137 | 2 I38876 | MHC class I antige |
| 18 | 19 | 76.0 | 137 | 2 I38860 | MHC class I antige |
| 19 | 19 | 76.0 | 137 | 2 I38874 | MHC class I antige |
| 20 | 19 | 76.0 | 137 | 2 E63559 | hypothetical prote |
| 21 | 19 | 76.0 | 138 | 2 F69130 | phosphoribosyl-AMP |
| 22 | 19 | 76.0 | 144 | 2 B69696 | ribosomal protein |
| 23 | 19 | 76.0 | 146 | 2 F64000 | hypothetical prote |
| 24 | 19 | 76.0 | 146 | 2 S40999 | hypothetical prote |
| 25 | 19 | 76.0 | 147 | 2 S75758 | hypothetical prote |
| 26 | 19 | 76.0 | 151 | 2 C75157 | protein tyrosine p |
| 27 | 19 | 76.0 | 152 | 2 A37815 | globin (domain E) |
| 28 | 19 | 76.0 | 164 | 2 B70760 | hypothetical prote |
| 29 | 19 | 76.0 | 165 | 2 F72600 | hypothetical prote |
| 30 | 19 | 76.0 | 181 | 2 I59188 | MHC cell surface g |

31 19 76.0 186 2 E65105 hypothetical 20.3
32 19 76.0 187 2 H69706 RNA polymerase ECF
33 19 76.0 191 2 D69867 hypothetical prote
34 19 76.0 192 2 I50123 RAG-1 protein - go
35 19 76.0 192 2 H70375 conserved hypothet
36 19 76.0 195 2 E27733 nifQ protein - Azo
37 19 76.0 197 2 G69459 hypothetical prote
38 19 76.0 202 2 G70752 probable regulator
39 19 76.0 203 1 BVECRV riva protein - Esc
40 19 76.0 205 2 C36795 hypothetical prote
41 19 76.0 205 2 T02685 hypothetical prote
42 19 76.0 208 2 E64753 yagK protein - Esc
43 19 76.0 208 2 A56912 FAD protein - hum
44 19 76.0 208 2 I38041 receptor-induced t
45 19 76.0 209 2 C70652 probable transcrip

ALIGNMENTS

RESULT 1
D69330
hypothetical protein AF0644 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-Jun-1998
C:Accession: D69330
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes,
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A:Reference number: A69250; MUID:98049343
A:Accession: D69330
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-35 <KLE>
A:Cross-references: GB:AE001060; GB:AE000782; NID:g2689383; PID:g2649986; TIGR:AF0644

Query Match 76.0%; Score 19; DB 2; Length 35;
Best Local Similarity 80.0%; Pred. NO. 49;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
DB 29 RETLR 33

RESULT 2
S01800
pepsin (EC 3.4.23.-) 3 precursor - North Pacific bluefin tuna (fragment)
C:Species: Thunnus thynnus orientalis (North Pacific bluefin tuna)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 08-Nov-1996
C:Accession: S01800
R:Tanji, M.; Kageyama, T.; Takahashi, K.
Eur. J. Biochem. 177, 251-259, 1988
A:Title: Tuna pepsinogens and pepsins. Purification, characterization and amino-termi
A:Reference number: S01798; MUID:89053692
A:Accession: S01800
A:Molecule type: protein
A:Residues: 1-60 <TAN>
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; protein digestion
F:1-35/Domain: propeptide #status experimental <PRO>
F:36-60/Product: pepsin (fragment) #status experimental <MAT>

Query Match 76.0%; Score 19; DB 2; Length 60;
Best Local Similarity 80.0%; Pred. NO. 85;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|| ||
DB 12 RESLR 16

RESULT 3

E70644
probable ribosomal protein L30 rpmD - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 24-Sep-1999
C:Accession: E70644
R:Colle, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, J.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987
A:Accession: E70644
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-65 <COL>
A:Cross-references: GB:284395; GB:AL123456; NID:g3361698; PIDN:CAB06446.1; PID:e293132;
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: rpmD
C:Superfamily: Escherichia coli ribosomal protein L30

Query Match 76.0%; Score 19; DB 2; Length 65;
Best Local Similarity 80.0%; Pred. No. 93;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|| ||
DB 20 RESLR 24

RESULT 4

S44892
Zkl12.4 protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 09-Sep-1997
C:Accession: S44892
R:Du, Z.
submitted to the EMBL Data Library, May 1993
A:Description: Sequence of the C. elegans cosmid zkl12.
A:Reference number: S44613
A:Accession: S44892
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <DUZ>
A:Cross-references: EMBL:L14324; NID:g289740; PID:g289747

Query Match 76.0%; Score 19; DB 2; Length 100;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|| ||
DB 69 RESLR 73

RESULT 5

S19733
hypothetical protein 2 - Thiobacillus versutus (fragment)
C:Species: Thiobacillus versutus
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Sep-1997
C:Accession: S19733
R:Ubbink, M.; van Kleef, M.A.G.; Kleinjan, D.J.; Hoitink, C.W.G.; Huitema, F.; Beintema, Eur. J. Biochem. 202, 1003-1012, 1991
A:Title: Cloning, sequencing and expression studies of the genes encoding amicyanin and

A:Reference number: S19730; MUID:92111471
A:Accession: S19733
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-102 <UBB>
A:Cross-references: GB:M58001; NID:gl54632; PID:gl54636

Query Match 76.0%; Score 19; DB 2; Length 102;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|| ||
DB 17 RESLR 21

RESULT 6

C75046
hypothetical protein PAB0900 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: C75046
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome s
A:Reference number: A75001
A:Accession: C75046
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-104 <KAW>
A:Cross-references: GB:AJ248287; GB:AL096836; NID:g5458657; PIDN:CAB50264.1; PID:e151
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB0900

Query Match 76.0%; Score 19; DB 2; Length 104;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|| ||
DB 95 RESLR 99

RESULT 7

S26322
Ig heavy chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1998 #sequence_revision 19-Mar-1998 #text_change 17-Apr-1998
C:Accession: S26322
R:Stark, S.E.; Caton, A.J.
J. Exp. Med. 174, 613-624, 1991
A:Title: Antibodies that are specific for a single amino acid interchange in a protein
A:Reference number: S26309; MUID:91341421
A:Accession: S26322
A:Molecule type: mRNA
A:Residues: 1-106 <STA>
A:Cross-references: EMBL:X59182
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin

Query Match 76.0%; Score 19; DB 2; Length 106;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
|| ||
DB 91 RESLR 95

RESULT 8

B44915

hoxL protein - Azotobacter vinelandii

C:Species: Azotobacter vinelandii

C:Date: 01-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 13-Sep-1998

C:Accession: B44915

R:Menon, A.L.; Mortenson, L.E.; Robson, R.L.

J. Bacteriol. 174, 4549-4557, 1992

A:Title: Nucleotide sequences and genetic analysis of hydrogen oxidation (hox) genes in

A:Reference number: A44915; MUID:92325046

A:Accession: B44915

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-106 <MEN>

A:Note: sequence extracted from NCBI backbone (NCBIN:108128, NCBIIP:108130)

C:Superfamily: hydrogenase related protein

Query Match

Best Local Similarity 76.0%; Score 19; DB 2; Length 106;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 REXLR 5

|| ||

Db 62 REALR 66

RESULT 9

S74019

hypothetical protein c0201/c0634 - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 17-Mar-1999

C:Accession: S74019

R:Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.

Mol. Microbiol. 22, 175-191, 1996

A:Title: Organizational characteristics and information content of an archaeal genome: 1

A:Reference number: S73076; MUID:97055432

A:Accession: S74019

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-106 <SEN>

A:Cross-references: EMBL:Y08256; NID:g1707679; PID:e284023; PID:g1707713

A:Experimental source: strain P2

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996

Query Match

Best Local Similarity 76.0%; Score 19; DB 2; Length 106;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 REXLR 5

|| ||

Db 14 REALR 18

RESULT 10

C24733

myosin alpha heavy chain, cardiac muscle - mouse (fragments)

C:Species: Mus musculus (house mouse)

C:Date: 24-Jun-1987 #sequence_revision 24-Jun-1987 #text_change 13-Feb-1998

C:Accession: C24733

R:Weydert, A.; Daubas, P.; Lazaridis, I.; Barton, P.; Garner, I.; Leader, D.P.; Bonhomme

Proc. Natl. Acad. Sci. U.S.A. 82, 7183-7187, 1985

A:Reference number: A94066; MUID:86042630

A:Accession: C24733

A:Molecule type: mRNA

A:Residues: 1-124 <WEY>

C:Superfamily: myosin heavy chain; myosin motor domain homology

C:Keywords: ATP; cardiac muscle; heart; muscle

Query Match

Best Local Similarity 76.0%; Score 19; DB 2; Length 124;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 REXLR 5

|| ||

Db 11 REALR 15

RESULT 11

C72247

ribosomal protein S11 - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 23-Jul-1999

C:Accession: C72247

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome

A:Reference number: A72200; MUID:99287316

A:Accession: C72247

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-130 <ARN>

A:Cross-references: GB:AE001798; GB:AE000512; NID:g4982033; PID:g4982040; TIGR:TM1474

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM1474

C:Superfamily: Escherichia coli ribosomal protein S11

Query Match

Best Local Similarity 76.0%; Score 19; DB 2; Length 130;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 REXLR 5

|| ||

Db 73 REALR 77

RESULT 12

JQ0370

photosystem I llk protein precursor - Chlamydomonas reinhardtii

N:Alternate names: photosystem I polypeptide 28; thylakoid polypeptide 28

C:Species: Chlamydomonas reinhardtii

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 26-Aug-1999

C:Accession: JQ0370; S06063

R:Franzen, L.G.; Frank, G.; Zuber, H.; Rochaix, J.D.

Mol. Gen. Genet. 219, 137-144, 1989

A:Title: Isolation and characterization of cDNA clones encoding photosystem I subunit

A:Reference number: JQ0370; MUID:90136501

A:Accession: JQ0370

A:Molecule type: mRNA

A:Residues: 1-130 <FRA>

A:Cross-references: EMBL:X15164; NID:g18191; PID:CAA33256.1; PID:g18192

C:Comment: Photosystem I catalyzes the light-driven electron transfer from plastocyan

C:Comment: This protein is an extrinsic membrane protein.

C:Superfamily: photosystem I protein psbH

C:Keywords: chloroplast; photosynthesis; photosystem I

F:1-30/Domain: transit peptide (chloroplast) #status predicted <TNP>

F:31-130/Product: photosystem I llk protein #status predicted <MAT>

Query Match

Best Local Similarity 76.0%; Score 19; DB 2; Length 130;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 REXLR 5

|| ||

Db 80 REALR 84

RESULT 13

S78260

ribosomal protein L18, chloroplast - Odontella sinensis chloroplast
C:Species: chloroplast Odontella sinensis
C:Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 13-Aug-1999
C:Accession: S78260
R:Kowallik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.
Plant Mol. Biol. Rep. 13, 336-342, 1995
A:Title: The Chloroplast Genome of a Chlorophyll a+c- containing Alga, Odontella sinensis
A:Reference number: S78238
A:Accession: S78260
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-135 <ROW>
A:Cross-references: EMBL:Z67753; NID:g1185127; PIDN:CAA91633.1; PID:e211826; PID:g118515
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
C:Genetics:
A:Gene: rpl18
A:Genome: Chloroplast
C:Superfamily: Escherichia coli ribosomal protein L18
C:Keywords: Chloroplast; protein biosynthesis; ribosome

Query Match 76.0%; Score 19; DB 2; Length 135;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5

Db 28 RESLR 32

RESULT 14

R5YM16

ribosomal protein L16 - Mycoplasma capricolum (SGC3)
N:Alternate names: protein MC008
C:Species: Mycoplasma capricolum
C:Date: 28-May-1986 #sequence_revision 31-Dec-1990 #text_change 22-Jun-1999
C:Accession: S02838; S77860; A02797; S48576
R:Ohkubo, S.; Muto, A.; Kawauchi, Y.; Yamao, F.; Osawa, S.
Mol. Gen. Genet. 210, 314-322, 1987
A:Title: The ribosomal protein gene cluster of Mycoplasma capricolum.
A:Reference number: S02838; MUID:88142549
A:Accession: S02838
A:Molecule type: DNA
A:Residues: 1-137 <OHK>
A:Cross-references: EMBL:X06414; NID:g44207; PIDN:CAA29711.1; PID:g44216
A:Experimental source: ATCC 27343
R:Bork, P.; Ouzounis, C.; Casari, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbert, W.;
Mol. Microbiol. 16, 955-967, 1995
A:Title: Exploring the Mycoplasma capricolum genome: a minimal cell reveals its physiology
A:Reference number: S77739; MUID:96059641
A:Accession: S77860
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-137 <BOR>
A:Cross-references: EMBL:Z33011; NID:g541684; PIDN:CAA83693.1; PID:g950060
A:Experimental source: ATCC 27343
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Genetics:
A:Gene: rpl16
A:Genetic code: SGC3
C:Superfamily: Escherichia coli ribosomal protein L16
C:Keywords: protein biosynthesis; ribosome

Query Match 76.0%; Score 19; DB 1; Length 137;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5

Db 115 REALR 119

RESULT 15
I38875
MHC class I antigen - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
C:Accession: I38875
R:Garber, T.L.; Butler, L.M.; Trachtenberg, E.A.; Erlich, H.A.; Rickards, O.; De Stef
Immunogenetics 42, 19-27, 1995
A:Title: HLA-B alleles of the Cayapa of Ecuador: new B39 and B15 alleles.
A:Reference number: I38860; MUID:95317819
A:Accession: I38875
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-137 <RES>
A:Cross-references: EMBL:U15639; NID:g930332; PIDN:AAA74046.1; PID:g930333
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 76.0%; Score 19; DB 2; Length 137;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5

Db 40 RESLR 44

Search completed: February 7, 2000, 11:54:13
Job time: 24323 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 00:59:42 ; Search time 63.71 Seconds
(without alignments)
4.688 Million cell updates/sec

Title: US-08-653-294-3
Perfect score: 25
Sequence: 1 REXLRXXXX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_38.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 19 | 76.0 | 60 | 1 PEP3_THUTO | P20141 thunnus thy |
| 2 | 19 | 76.0 | 65 | 1 RL30_MYCTU | P95070 mycobacteri |
| 3 | 19 | 76.0 | 71 | 1 RL30_MYCLE | O33001 mycobacteri |
| 4 | 19 | 76.0 | 100 | 1 YOG4_CAEEL | P34613 caenorhabdi |
| 5 | 19 | 76.0 | 106 | 1 HOXL_AZOVI | P40592 azotobacter |
| 6 | 19 | 76.0 | 130 | 1 PSAH_CHLRE | P13352 chlamydomon |
| 7 | 19 | 76.0 | 135 | 1 RX18_ODOSI | P49554 odontella s |
| 8 | 19 | 76.0 | 137 | 1 RL16_MYCCA | P02415 mycoplasma |
| 9 | 19 | 76.0 | 137 | 1 RL16_SPICI | O31162 spiroplasma |
| 10 | 19 | 76.0 | 138 | 1 HIS3_METTH | O26347 methanobact |
| 11 | 19 | 76.0 | 144 | 1 RL16_BACSU | P14577 bacillus su |
| 12 | 19 | 76.0 | 146 | 1 Y074_HAEIN | P43934 haemophilus |
| 13 | 19 | 76.0 | 152 | 1 GEB7_ARTSX | P19364 artemia sp. |
| 14 | 19 | 76.0 | 156 | 1 SOXR_PSEAE | Q51506 pseudomonas |
| 15 | 19 | 76.0 | 164 | 1 YK12_MYCTU | Q10845 mycobacteri |
| 16 | 19 | 76.0 | 186 | 1 YHBO_ECOLI | P45470 escherichia |
| 17 | 19 | 76.0 | 187 | 1 SIGW_BACSU | Q45585 bacillus su |
| 18 | 19 | 76.0 | 195 | 1 NIFQ_AZOVI | P11068 azotobacter |
| 19 | 19 | 76.0 | 202 | 1 YC55_MYCTU | Q11063 mycobacteri |
| 20 | 19 | 76.0 | 203 | 1 RUVA_ECOLI | P08576 escherichia |
| 21 | 19 | 76.0 | 205 | 1 VG02_HSVBP | P28979 equine herp |
| 22 | 19 | 76.0 | 208 | 1 FADD_HUMAN | Q13158 homo sapien |
| 23 | 19 | 76.0 | 208 | 1 YAGK_ECOLI | P77657 escherichia |
| 24 | 19 | 76.0 | 212 | 1 IFE4_CAEEL | Q22888 caenorhabdi |
| 25 | 19 | 76.0 | 214 | 1 ER3_MOUSE | P48299 mus musculu |
| 26 | 19 | 76.0 | 241 | 1 YMFC_BACSU | O31761 bacillus su |
| 27 | 19 | 76.0 | 254 | 1 RNC_TREPA | Q83787 treponema p |
| 28 | 19 | 76.0 | 261 | 1 YD53_MYCTU | O11023 mycobacteri |
| 29 | 19 | 76.0 | 270 | 1 LB33_HUMAN | P01890 homo sapien |
| 30 | 19 | 76.0 | 274 | 1 UCRI_BOVIN | P13272 bos taurus |
| 31 | 19 | 76.0 | 292 | 1 CGD3_HUMAN | P30281 homo sapien |
| 32 | 19 | 76.0 | 292 | 1 CGD3_MOUSE | P30282 mus musculu |
| 33 | 19 | 76.0 | 292 | 1 MAUJ_PARVE | Q56464 paracoccus |
| 34 | 19 | 76.0 | 293 | 1 CGD3_RAT | P48961 rattus norv |

35 19 76.0 299 1 YHJC_ECOLI P37641 escherichia
36 19 76.0 308 1 NOSF_PSEST P19844 pseudomonas
37 19 76.0 310 1 LDH_THECA P06150 thermus aqu
38 19 76.0 329 1 DHQA_EMENI P25415 emericella
39 19 76.0 329 1 HEM2_MYCLE P46723 mycobacteri
40 19 76.0 330 1 HEM2_STRCO P4919 streptomyce
41 19 76.0 334 1 BIOB_CORGL P46396 corynebacte
42 19 76.0 340 1 THPS_SULAC P17118 sulfolobus
43 19 76.0 354 1 IB24_HUMAN P30470 homo sapien
44 19 76.0 354 1 CD72_MOUSE P21855 mus musculu
45 19 76.0 356 1 CFXU_BRAJA Q59203 bradyrhizob

ALIGNMENTS

RESULT 1
PEP3_THUTO
ID PEP3_THUTO STANDARD; PRT; 60 AA.
AC P20141;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE PEP3THUN3 (EC 3.4.23.-) (FRAGMENT).
OS Thunnus thynnus orientalis (North Pacific bluefin tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Perciformes; Scombroidei; Scombridae; Thunnus.
RN [1]
RP SEQUENCE.
RX MEDLINE; 89052692.
RA TANJ M., KAGEYAMA T., TAKAHASHI K.;
RT "Tuna pepsinogens and pepsins. Purification, characterization and
RT amino-terminal sequences".
RL Eur. J. Biochem. 177:251-259(1988).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
DR PIR; S01800; S01800.
DR HSSP; P56272; 1AM5.
DR PROSITE; PS00141; ASP_PROTEASE; PARTIAL.
KW Hydrolase; Aspartyl protease; Zymogen.
FT PROPEP 1 35 ACTIVATION PEPTIDE.
FT CHAIN 36 >60 PEP3IN 3.
FT NON_TER 60 60
SQ SEQUENCE 60 AA; 6970 MW; B1AD33F1 CRC32;
Query Match 76.0%; Score 19; DB 1; Length 60;
Best Local Similarity 80.0%; Pred. No. 47;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 REXLR 5
Db 12 RESLR 16

RESULT 2
RL30_MYCTU
ID RL30_MYCTU STANDARD; PRT; 65 AA.
AC P95070;
DT 15-DEC-1999 (Rel. 39, Created)
DT 15-DEC-1999 (Rel. 39, Last sequence update)
DT 15-DEC-1999 (Rel. 39, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L30.
GN RPSD OR RV0722 OR MFCY210.41.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE; 98295987.
RA COLE S.T., BROSC R., PARKHILL J., GARNIER T., CHURCHER C., HARRIS D.,

RA GORDON S.V., EIGLMEIER K., GAS S., BARRY C.E. III, TEKAIA F.,
 RA BADCOCK K., BASHAM D., BROWN D., CHILLINGWORTH T., CONNOR R.,
 RA DAVIES R., DEVLIN K., FELTWELL T., GENTILES S., HAMLIN N., HOLROYD S.,
 RA HORNBRY T., JAGELS K., KROGH A., MCLEAN J., MOULE S., MURPHY L.,
 RA OLIVER S., OSBORNE J., QUAIL M.A., RAJANDREAM M.A., ROGERS J.,
 RA RUTTER S., SEEGER K., SKELTON S., SQUARES S., SQUARES R., SULSTON J.E.,
 RA TAYLOR K., WHITEHEAD S., BARRELL B.G.,
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.",
 RL Nature 393:537-544(1998).
 CC -1- SIMILARITY: BELONGS TO THE L30P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z84395; CAB06446.1; -
 DR PROSITE; PS00634; RIBOSOMAL_L30; FALSE_NEG.
 DR PFAM; PF00327; Ribosomal_L30; 1.
 KW Ribosomal protein.
 SQ SEQUENCE 65 AA; 7346 MW; 9253D057 CRC32;

Query Match 76.0%; Score 19; DB 1; Length 65;
 Best Local Similarity 80.0%; Pred. No. 51;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 || ||
 Db 20 RESLR 24

RESULT 3
 RL30_MYCLE
 ID RL30_MYCLE STANDARD; PRT; 71 AA.
 AC O33001;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE 50S RIBOSOMAL PROTEIN L30.
 GN RPMO OR MCB2492.22.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA COLE S.T., FLESELLES B., HONORE N.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE L30P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z98756; CAB11454.1; -
 DR PROSITE; PS00634; RIBOSOMAL_L30; FALSE_NEG.
 DR PFAM; PF00327; Ribosomal_L30; 1.
 KW Ribosomal protein.
 SQ SEQUENCE 71 AA; 7886 MW; 779051E0 CRC32;

Query Match 76.0%; Score 19; DB 1; Length 71;
 Best Local Similarity 80.0%; Pred. No. 56;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 || ||
 Db 20 RESLR 24

RESULT 4
 YOG4_CAEEL
 ID YOG4_CAEEL STANDARD; PRT; 100 AA.
 AC F34613;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE HYPOTHETICAL 11.2 KD PROTEIN ZK112.4 IN CHROMOSOME III.
 GN ZK112.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE; 94150718
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FRASER A.,
 RA FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,
 RA JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N.,
 RA LATREILLE P., LIGHTNING J., LLOYD C., MORTIMORE B., O'CALLAGHAN M.,
 RA FARNSON J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SIMS M., SMALDON N., SMITH A., SMITH M., SONNHAMMER E., STADEN R.,
 RA SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K.,
 RA WATSON R., WATSON A., WEINSTOCK L., WILKINSON-SPROAT J.,
 RA WOHLDMAN P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.",
 RL Nature 368:32-38(1994).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L14324; AAA28187.1; -
 DR PIR; S44892; S44892.
 DR WORMPEP; ZK112.4; CE00375.
 KW Hypothetical protein.
 SQ SEQUENCE 100 AA; 11248 MW; 52F18207 CRC32;

Query Match 76.0%; Score 19; DB 1; Length 100;
 Best Local Similarity 80.0%; Pred. No. 81;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 || ||
 Db 69 REALR 73

RESULT 5
 HOXL_AZOVI
 ID HOXL_AZOVI STANDARD; PRT; 106 AA.
 AC F40592;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE HYDROGENASE EXPRESSION/FORMATION PROTEIN HOXL.
 GN HOXL.
 OS Azotobacter vinelandii.
 OC Bacteria; Proteobacteria; gamma subdivision; Azotobacteraceae;
 OC Azotobacter.
 RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-UW.
RX MEDLINE: 92325046.
RA MENON A., MORTENSON L.E., ROBSON R.L.;
RT "Nucleotide sequences and genetic analysis of hydrogen oxidation
RT (hox) genes in Azotobacter vinelandii.";
RL J. Bacteriol. 174:4549-4557(1992).
CC -!- SIMILARITY: BELONGS TO THE HUPF/HYPC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M80522; AAA22127.1; -.
DR EMBL: L23970; AAA19502.1; -.
DR PIR: B44915; B44915.
DR PROSITE: PS01097; HUPF_HYPC; 1.
DR PIR: PF01455; Hupf_Hypc; 1.
SQ SEQUENCE 106 AA; 11436 MW; 3113532D CRC32;

Query Match 76.0%; Score 19; DB 1; Length 106;
Best Local Similarity 80.0%; Pred. No. 86;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 REXLR 5
DB 62 REALR 66

RESULT 6
ID PSAH_CHLRE STANDARD; PRT; 130 AA.
AC P13352;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE PHOTOSYSTEM I REACTION CENTRE SUBUNIT VI PRECURSOR (LIGHT-HARVESTING
DE COMPLEX I 11 KD PROTEIN) (PSI-H) (P28 PROTEIN).
GN PSAH.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-137C;
RX MEDLINE: 90136501.
RA FRANZEN L.-G., FRANK G., ZUBER H., ROCHAIX J.-D.;
RT "Isolation and characterization of cDNA clones encoding photosystem I
RT subunits with molecular masses 11.0, 10.0 and 8.4 kDa from
RT Chlamydomonas reinhardtii.";
RL Mol. Gen. Genet. 219:137-144(1989).
CC -!- FUNCTION: POSSIBLE ROLE COULD BE THE DOCKING OF THE LHC I ANTENNA
CC COMPLEX TO THE CORE COMPLEX.
CC -!- SIMILARITY: BELONGS TO THE PSAH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15164; CAA33256.1; -.
DR PIR: S06063; S06063.
DR PIR: JQ0370; JQ0370.
DR MENDEL: 7698; CHLre:PSAH;1.
KW Photosynthesis; Photosystem I; Chloroplast; Transit peptide.

FT TRANSIT 1 30 CHLOROPLAST.
FT CHAIN 31 130 PHOTOSYSTEM I REACTION CENTRE SUBUNIT VI.
SQ SEQUENCE 130 AA; 14151 MW; 37AE862 CRC32;

Query Match 76.0%; Score 19; DB 1; Length 130;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 REXLR 5
DB 80 RESLR 84

RESULT 7
ID RK18_ODOSI STANDARD; PRT; 135 AA.
AC P49554;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L18.
GN RPL18.
OS Odontella sinensis.
OC Chloroplast.
OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Biddulphiophycidae; Eupodiscales; Eupodiscaeae; Odontellia.
RN [1]
RP SEQUENCE FROM N.A.
RA KOWALLIK K.V., STOEBE B., SCHAFFRAN I., KROTH-PANCIC P., FREIER U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
RT Odontella sinensis.";
RL Plant Mol. Biol. Rep. 13:336-342(1995).
CC -!- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z67753; CAA91633.1; -.
DR MENDEL: 9483; ODOSI:rpl18;1.
DR PIR: PF00861; Ribosomal_L18p; 1.
SQ SEQUENCE 135 AA; 15417 MW; 524017AC CRC32;

Query Match 76.0%; Score 19; DB 1; Length 135;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 REXLR 5
DB 28 RESLR 32

RESULT 8
ID RL16_MYCCA STANDARD; PRT; 137 AA.
AC P02415;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L16.
GN RPLP.
OS Mycoplasma capricolum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC capricolum group.
RN [1]
RP SEQUENCE FROM N.A.

RX STRAIN-ATCC 27343 / KID;
 RA OHKUBO S., MUTO A., KAWAUCHI Y., YAMAO F., OSAWA S.;
 RT "The ribosomal protein gene cluster of *Mycoplasma capricolum*.";
 RL Mol. Gen. Genet. 210:314-322(1987).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-ATCC 27343 / KID;
 RX MEDLINE; 96059641.
 RA BORK P., OUZOUNIS C., CASARI G., SCHNEIDER R., SANDER C., DOLAN M.,
 RA GILBERT W., GILBERT P.M.;
 RT "Exploring the *Mycoplasma capricolum* genome: a minimal cell reveals
 RT its physiology.";
 RL Mol. Microbiol. 16:955-967(1995).
 RN [3]
 RP SEQUENCE OF 1-117 FROM N.A.
 RX MEDLINE; 85190486.
 RA YAMAO F., MUTO A., KAWAUCHI Y., IWAMI M., IWAGAMI S., AZUMI Y.,
 RA OSAWA S.;
 RT "UGA is read as tryptophan in *Mycoplasma capricolum*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2306-2309(1985).
 CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS
 CC LOCATED AT THE A SITE OF THE PEPTIDYLTRANSFERASE CENTER
 CC (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE L16P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X06414; CAA29711.1; -;
 DR EMBL; Z33011; CAA83693.1; -;
 DR EMBL; K02973; AAA25439.1; -;
 DR PIR; S02838; R5YM16.
 DR PROSITE; PS00586; RIBOSOMAL_L16_1; 1.
 DR PROSITE; PS00701; RIBOSOMAL_L16_2; 1.
 DR PFAM; PF00252; Ribosomal_L16; 1.
 KW Ribosomal protein; rRNA-binding.
 SQ SEQUENCE 137 AA; 15799 MW; 68B1D725 CRC32;

 Query Match 76.08; Score 19; DB 1; Length 137;
 Best Local Similarity 80.08; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 1 RECLR 5
 DB 115 RECLR 119

 RESULT 9
 RL16 SPIC1
 ID RL16 SPIC1 STANDARD; PRT; 137 AA.
 AC Q31162;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE 50S RIBOSOMAL PROTEIN L16.
 GN RPLP.
 OS *Spiroplasma citri*.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Spiroplasmataceae; Spiroplasma.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 27556 / R8A2;
 RA LE DANTEC L., SAILLARD C., BOVE J.M.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS
 CC LOCATED AT THE A SITE OF THE PEPTIDYLTRANSFERASE CENTER

CC (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE L16P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF031160; AAC35871.1; -;
 DR EMBL; PS00586; RIBOSOMAL_L16_1; 1.
 DR PROSITE; PS00701; RIBOSOMAL_L16_2; 1.
 DR PFAM; PF00252; Ribosomal_L16; 1.
 KW Ribosomal protein; rRNA-binding.
 SQ SEQUENCE 137 AA; 15742 MW; 971D639D CRC32;

 Query Match 76.08; Score 19; DB 1; Length 137;
 Best Local Similarity 80.08; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 1 RECLR 5
 DB 115 RECLR 119

 RESULT 10
 HIS3 METH
 ID HIS3 METH STANDARD; PRT; 138 AA.
 AC Q26347;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DE PHOSPHORIBOSYL-AMP CYCLOHYDROLASE (EC 3.5.4.19).
 GN HIS1 OR MTH245.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 OC Methanobacterium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DELTA H;
 RX MEDLINE; 98037514.
 RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,
 RA ALDRIDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,
 RA HARRISON D., HOANG L., KEAGLE P., LOMM W., POTTHIER B., QIU D.,
 RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,
 RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,
 RA MCDUGALL S., SHIMER G., GOVAL A., PIETROVSKI S., CHURCH G.M.,
 RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;
 RT "Complete genome sequence of *Methanobacterium thermoautotrophicum*
 RT DeltaH: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -!- CATALYTIC ACTIVITY: 5-PHOSPHORIBOSYL-AMP + H(2)O =
 CC 5-(5-PHOSPHO-D-RIBOSYLAMINOFORMIMINO)-1-(5-PHOSPHO-RIBOSYL)
 CC IMIDAZOLE-4-CARBOXAMIDE
 CC -!- PATHWAY: THIRD STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE000811; AAB84751.1; -;
 DR PFAM; PF01502; PRA-CH; 1.
 KW Histidine biosynthesis; Hydrolase.
 SQ SEQUENCE 138 AA; 15458 MW; BAEB24D6 CRC32;

Query Match 76.0%; Score 19; DB 1; Length 138;
 Best Local Similarity 80.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 REXLR 5
 || ||
 Db 44 REALR 48

RESULT 11
 RL16_BACSU STANDARD; PRT: 144 AA.
 ID RL16_BACSU
 AC P14577;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE 50S RIBOSOMAL PROTEIN L16.
 GN RPLP.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 RN [1]
 RP SEQUENCE OF 1-55 FROM N.A.
 RC STRAIN=168;
 RA YOSHIKAWA H., YASUMOTO K., TAKAHASHI H.;
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SG38;
 RA LI X., LINDAHL L., ZENGEL J.M.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 39-144 FROM N.A.
 RC STRAIN=168;
 RX MEDLINE; 90016806.
 RA HENKIN T.M., MOON S.H., MATTHEAKIS L.C., NOMURA M.;
 RT "Cloning and analysis of the spc ribosomal protein operon of Bacillus
 RT subtilis: comparison with the spc operon of Escherichia coli.";
 RL Nucleic Acids Res. 17:7469-7486(1989).
 RN [4]
 RP SEQUENCE OF 96-144 FROM N.A.
 RC STRAIN=168 / MAREBURG;
 RX MEDLINE; 96186897.
 RA SUH J.W., BOYLAN S.A., OH S.H., PRICE C.W.;
 RT "Genetic and transcriptional organization of the Bacillus subtilis
 RT spc-alpha region.";
 RL Gene 169:17-23(1996).
 CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS
 CC LOCATED AT THE A SITE OF THE PEPTIDYLTRANSFERASE CENTER.
 CC -!- SIMILARITY: BELONGS TO THE L16P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; D50302; BAA0838.1; -;
 DR EMBL; D50303; -; NOT ANNOTATED_CDS.
 DR EMBL; U43929; AAC45963.1; -;
 DR EMBL; X15664; CAA33698.1; -;
 DR EMBL; L47971; AAB06806.1; -;
 DR EMBL; Z99104; CAB11899.1; -;
 DR PIR; S05989; S05989.
 DR SUBTILIST; BG10755; RPLP.
 DR PROSITE; PS00586; RIBOSOMAL_L16_1; 1.
 DR PROSITE; PS00701; RIBOSOMAL_L16_2; 1.
 DR PFAM; PF00252; Ribosomal_L16; 1.
 KW Ribosomal protein; rRNA-binding.
 SQ SEQUENCE 144 AA; 16190 MW; 8580A491 CRC32;

Query Match 76.0%; Score 19; DB 1; Length 144;
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 REXLR 5
 || ||
 Db 115 REALR 119

RESULT 12
 Y074_HAEIN STANDARD; PRT: 146 AA.
 ID Y074_HAEIN
 AC P43934;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE HYPOTHETICAL PROTEIN HI0074.
 GN HI0074.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RD / KW20;
 RX MEDLINE; 95350630.
 RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
 RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,
 RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
 RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,
 RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,
 RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
 RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEORHAGEN N.S.M.,
 RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
 RA VENTER J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus
 RT influenzae Rd.";
 RL Science 269:496-512(1995).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U32693; AAC21759.1; -;
 DR TIGR; HI0074; -;
 KW Hypothetical protein.
 SQ SEQUENCE 146 AA; 17127 MW; A139C9F4 CRC32;

Query Match 76.0%; Score 19; DB 1; Length 146;
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 REXLR 5
 || ||
 Db 85 REALR 89

RESULT 13
 GLB7_ARTSX STANDARD; PRT: 152 AA.
 ID GLB7_ARTSX
 AC P19364;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1990 (Rel. 16, Last annotation update)
 DE GLOBIN E7, EXTRACELLULAR.
 OS Artemia sp. (Brine shrimp).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;

```

OC Artemiidae; Artemia.
RN [1]
RP SEQUENCE.
RX MEDLINE; 90354411.
RA MOENS L., VAN HAUWAERT M.-L., DE SMET K., VER DONCK K.,
RT "STRUCTURAL INTERPRETATION OF THE AMINO ACID SEQUENCE OF A SECOND
RT domain from the Artemia covalent polymer globin.";
RL J. Biol. Chem. 265:14285-14291(1990).
CC -!- SUBUNIT: ARTEMIA HEMOGLOBIN IS A DIMER OF TWO SIMILAR SIZED
CC SUBUNITS. EACH SUBUNIT REPRESENTS A GLOBIN CHAIN WHICH EXISTS IN
CC TWO FORMS (ALPHA AND BETA), THUS MAKING POSSIBLE THREE DIFFERENT
CC PHENOTYPES (HB1, ALPHA(2), HB2, ALPHA/BETA, HB3, BETA(2)). THE
CC GLOBIN CHAIN IS A POLYMER OF EIGHT HEME-BINDING COVALENTLY LINKED
CC DOMAINS.
DR PIR; A37815; A37815.
DR PROSITE; PS01033; GLOBIN; 1.
DR PFAM; PF00042; globin; 1.
KW Heme; Oxygen transport; Respiratory protein.
SQ SEQUENCE 152 AA; 17074 MW; 6058429B CRC32;

Query Match 76.0%; Score 19; DB 1; Length 152;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
DB 82 RESLR 86

RESULT 14
SOXR_PSEAE STANDARD; PRT; 156 AA.
AC Q51506;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SOXR PROTEIN.
GN SOXR.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;
OC Pseudomonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PAO1;
RA LIAO X., HANCOCK R.E.W.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATES THE TRANSCRIPTION OF THE SOXR GENE WHICH
CC ITSELF CONTROL THE SUPEROXIDE RESPONSE REGULON. SOXR CONTAINS A
CC 2FE-2S IRON-SULFUR CLUSTER THAT MAY ACT AS A REDOX SENSOR SYSTEM
CC THAT RECOGNIZES SUPEROXIDE. THE VARIABLE REDOX STATE OF THE FE-S
CC CLUSTER MAY THUS BE EMPLOYED IN VIVO TO MODULATE THE
CC TRANSCRIPTIONAL ACTIVITY OF SOXR IN RESPONSE TO SPECIFIC TYPES OF
CC OXIDATIVE STRESS (BY SIMILARITY).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE MERR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X95517; CAA64771.1; -
CC PROSITE; PS00552; HTH_MERR_FAMILY; 1.
DR PFAM; PF00376; merr; 1.
KW DNA-binding; Transcription regulation; Activator; Iron-sulfur.
FT DNA_BIND 12 31 H-T-H MOTIF (POTENTIAL).
FT DOMAIN 117 128 MIGHT BE PART OF A SENSOR REGION.

```

```

SQ SEQUENCE 156 AA; 16998 MW; 488642E3 CRC32;

Query Match 76.0%; Score 19; DB 1; Length 156;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
DB 49 RETLR 53

RESULT 15
YK12_MYCTU STANDARD; PRT; 164 AA.
ID YK12_MYCTU
AC Q10845;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1999 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 18.2 KD PROTEIN RV2012.
GN RV2012 OR MTCV39.05C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE; 98295987.
RA COLE S.T., BROSCHE R., PARKHILL J., GARNIER T., CHURCHER C., HARRIS D.,
RA GORDON S.V., EIGLMEIER K., GAS S., BARRY C.E. III, TERRAIA F.,
RA BADCOCK K., BASHAM D., BROWN D., CHILLINGWORTH T., CONNOR R.,
RA DAVIES T., DEVLIN K., FELTWEILL T., GENTLES S., HAMLIN N., HOLROYD S.,
RA HORNSBY T., JAGELS K., KROGH A., MCLEAN J., MOULE S., MURPHY L.,
RA OLIVER S., OSBORNE J., QUAIL M.A., RAJANDREAM M.A., ROGERS J.,
RA RUTHER S., SEEGER K., SHELTON S., SQUARES S., SQUARES R., SULSTON J.E.,
RA TAYLOR K., WHITEHEAD S., BARRELL B.G.;
RL "Deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence.";
RL Nature 393:537-544(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 274025; CAA98403.1; -
CC Hypothetical protein.
SQ SEQUENCE 164 AA; 18202 MW; 0073FA55 CRC32;

Query Match 76.0%; Score 19; DB 1; Length 164;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
DB 12 RETLR 16

Search completed: February 8, 2000, 00:59:44
Job time: 3773 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 13:17:26 ; Search time 209.03 Seconds
(without alignments)
3.317 Million cell updates/sec

Title: US-08-653-294-3
Perfect score: 25
Sequence: 1 REXLRXXXX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

SPTREMBL_12:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 19 | 76.0 | 35 | 1 029613 | 029613 archaeoglob |
| 2 | 19 | 76.0 | 47 | 7 019787 | 019787 homo sapien |
| 3 | 19 | 76.0 | 61 | 5 P91619 | P91619 drosophila |
| 4 | 19 | 76.0 | 81 | 7 019523 | 019523 homo sapien |
| 5 | 19 | 76.0 | 81 | 7 019525 | 019525 homo sapien |
| 6 | 19 | 76.0 | 81 | 7 019527 | 019527 homo sapien |
| 7 | 19 | 76.0 | 81 | 7 019529 | 019529 homo sapien |
| 8 | 19 | 76.0 | 81 | 7 019531 | 019531 homo sapien |
| 9 | 19 | 76.0 | 81 | 7 019533 | 019533 homo sapien |
| 10 | 19 | 76.0 | 81 | 7 019535 | 019535 homo sapien |
| 11 | 19 | 76.0 | 81 | 7 019537 | 019537 homo sapien |
| 12 | 19 | 76.0 | 83 | 7 019539 | 019539 homo sapien |
| 13 | 19 | 76.0 | 83 | 7 019541 | 019541 homo sapien |
| 14 | 19 | 76.0 | 83 | 7 019543 | 019543 homo sapien |
| 15 | 19 | 76.0 | 89 | 4 095956 | 095956 homo sapien |
| 16 | 19 | 76.0 | 89 | 4 095452 | 095452 homo sapien |
| 17 | 19 | 76.0 | 89 | 7 P9615 | P9615 homo sapien |
| 18 | 19 | 76.0 | 89 | 7 P79620 | P79620 homo sapien |
| 19 | 19 | 76.0 | 89 | 7 019548 | 019548 homo sapien |
| 20 | 19 | 76.0 | 89 | 7 019549 | 019549 homo sapien |

| | | | | | |
|----|----|------|-----|-----------|--------------------|
| 21 | 19 | 76.0 | 89 | 7 019550 | 019550 homo sapien |
| 22 | 19 | 76.0 | 89 | 7 019551 | 019551 homo sapien |
| 23 | 19 | 76.0 | 89 | 7 019640 | 019640 homo sapien |
| 24 | 19 | 76.0 | 89 | 7 077959 | 077959 homo sapien |
| 25 | 19 | 76.0 | 89 | 7 077967 | 077967 homo sapien |
| 26 | 19 | 76.0 | 89 | 7 078077 | 078077 homo sapien |
| 27 | 19 | 76.0 | 89 | 7 078174 | 078174 homo sapien |
| 28 | 19 | 76.0 | 89 | 7 019638 | 019638 homo sapien |
| 29 | 19 | 76.0 | 89 | 7 P79487 | P79487 homo sapien |
| 30 | 19 | 76.0 | 89 | 7 019587 | 019587 homo sapien |
| 31 | 19 | 76.0 | 89 | 7 019576 | 019576 homo sapien |
| 32 | 19 | 76.0 | 89 | 7 019577 | 019577 homo sapien |
| 33 | 19 | 76.0 | 89 | 7 019581 | 019581 homo sapien |
| 34 | 19 | 76.0 | 89 | 7 019583 | 019583 homo sapien |
| 35 | 19 | 76.0 | 90 | 2 086794 | 086794 streptomyce |
| 36 | 19 | 76.0 | 91 | 2 Q56376 | Q56376 escherichia |
| 37 | 19 | 76.0 | 96 | 5 P90836 | P90836 caenorhabdi |
| 38 | 19 | 76.0 | 96 | 6 077729 | 077729 ovis aries |
| 39 | 19 | 76.0 | 101 | 11 088517 | 088517 mus musculu |
| 40 | 19 | 76.0 | 106 | 1 P95881 | P95881 sulfolobus |
| 41 | 19 | 76.0 | 108 | 2 Q56375 | Q56375 escherichia |
| 42 | 19 | 76.0 | 109 | 2 087365 | 087365 staphylococ |
| 43 | 19 | 76.0 | 111 | 12 Q82684 | Q82684 infectious |
| 44 | 19 | 76.0 | 111 | 12 Q89837 | Q89837 infectious |
| 45 | 19 | 76.0 | 112 | 7 Q92671 | Q92671 homo sapien |

ALIGNMENTS

RESULT 1

O29613 ID O29613 PRELIMINARY: PRT: 35 AA.

AC O29613;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-AUG-1998 (Tremblrel. 07, Last annotation update)
DE HYPOTHETICAL 4.1 KD PROTEIN.

GN AF0644.

OS Archaeoglobus fulgidus.

OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;

OC Archaeoglobus.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;

RX MEDLINE; 98049343.

RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,

RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,

RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRPIDES N.C.,

RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,

RA KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,

RA PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,

RA OVERBEER K., GOCAYNE J.D., WEIDMAN J.F., McDONALD L., UTERBACK T.,

RA COTTON M.D., SPRIGGS T., ARTIACH P., KAINE B.P., SYKES S.M.,

RA SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,

RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,

RA VENTER J.C.;

RT "The complete genome sequence of the hyperthermophilic, sulphate-

reducing archaeon Archaeoglobus fulgidus.";

RL Nature 390:364-370(1997).

DR EMBL; AE001060; AAB90607.1; .

DR TIGR; AF0644; .

KW Hypothetical protein.

SO SEQUENCE 35 AA; 4056 MW; 9F03E370 CRC32;

Query Match

Best Local Similarity 80.0%; Score 19; DB 1; Length 35;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 REXLR 5

Db 29 RETLR 33

```

RESULT 2
O19787 ID O19787 PRELIMINARY; PRT; 47 AA.
AC O19787;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE HLA CLASS I ANTIGEN (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94090601.
RA YOSHIDA M., KIMURA A., KATSURAGI K., NUMANO F., SASAZUKI T.;
RT "DNA typing of HLA-B gene in Takayasu's arteritis.";
RL Tissue Antigens 42:87-90(1993).
DR EMBL; S67638; AAB29305.1; -
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 47
SQ SEQUENCE 47 AA; 5608 MW; E52E443A CRC32;

Query Match 76.0%; Score 19; DB 7; Length 47;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
DB 41 RESLR 45

RESULT 3
P91619 ID P91619 PRELIMINARY; PRT; 61 AA.
AC P91619;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)
DE FRU PROTEIN (FRAGMENT).
GN FRU.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-CANTON-S;
RX MEDLINE; 96382528.
RA ITO H., FUJITANI K., USUI K., SHIMIZU-NISHIKAWA K., TANAKA S.,
RA YAMAMOTO D.;
RT "Sexual orientation in Drosophila is altered by the satori mutation in
RT the sex-determination gene fruitless that encodes a zinc finger
RT protein with a BTB domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:9687-9692(1996).
DR EMBL; D84438; BAA12664.1; -
DR FLYBASE; FBgn0004652; fru.
FT NON_TER 61
FT NON_TER 61
SQ SEQUENCE 61 AA; 6954 MW; 301C8FA1 CRC32;

Query Match 76.0%; Score 19; DB 5; Length 61;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
DB 26 RESLR 30

RESULT 4
O19523 ID O19523 PRELIMINARY; PRT; 81 AA.
AC O19523;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
RA RUNGROUNG E., BEJCHANDRA S.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF014769; AAB67807.1; -
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 81
SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 76.0%; Score 19; DB 7; Length 81;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
DB 66 RESLR 70

RESULT 5
O19525 ID O19525 PRELIMINARY; PRT; 81 AA.
AC O19525;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
RA RUNGROUNG E., BEJCHANDRA S.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF014771; AAB67809.1; -
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 81
SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 76.0%; Score 19; DB 7; Length 81;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
DB 66 RESLR 70

RESULT 6
O19527 ID O19527 PRELIMINARY; PRT; 81 AA.

```

AC O19527;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
RA RUNGRONG E., BEJCHANDRA S.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF014773; AAB67811.1; -
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 81
SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 76.0%; Score 19; DB 7; Length 81;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
DB 66 RESLR 70

RESULT 7
O19529
ID O19529 PRELIMINARY; PRT; 81 AA.
AC O19529;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
RA RUNGRONG E., BEJCHANDRA S.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF014775; AAB67813.1; -
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 81
SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 76.0%; Score 19; DB 7; Length 81;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
DB 66 RESLR 70

RESULT 8
O19531
ID O19531 PRELIMINARY; PRT; 81 AA.
AC O19531;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).

GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
RA RUNGRONG E., BEJCHANDRA S.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF014777; AAB67815.1; -
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 81
SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 76.0%; Score 19; DB 7; Length 81;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
DB 66 RESLR 70

RESULT 9
O19533
ID O19533 PRELIMINARY; PRT; 81 AA.
AC O19533;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
RA RUNGRONG E., BEJCHANDRA S.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF014779; AAB67817.1; -
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 81
SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 76.0%; Score 19; DB 7; Length 81;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
DB 66 RESLR 70

RESULT 10
O19535
ID O19535 PRELIMINARY; PRT; 81 AA.
AC O19535;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]

RP SEQUENCE FROM N.A.
 RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
 RA RUNGRONG E., BEJCHANDRA S.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF014781; AAB67821.1; -
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 81 81
 SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 76.0%; Score 19; DB 7; Length 81;
 Best Local Similarity 80.0%; Pred. No. 3.4e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 || ||
 Db 66 RESLR 70

RESULT 11
 ID O19537 PRELIMINARY; PRT; 81 AA.
 AC O19537;
 DT 01-JAN-1998 (TREMELrel. 05, Created)
 DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)
 DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
 RA RUNGRONG E., BEJCHANDRA S.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF014783; AAB67821.1; -
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 81 81
 SQ SEQUENCE 81 AA; 9405 MW; 073087CE CRC32;

Query Match 76.0%; Score 19; DB 7; Length 81;
 Best Local Similarity 80.0%; Pred. No. 3.4e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 || ||
 Db 66 RESLR 70

RESULT 12
 ID O19539 PRELIMINARY; PRT; 83 AA.
 AC O19539;
 DT 01-JAN-1998 (TREMELrel. 05, Created)
 DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)
 DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
 RA RUNGRONG E., BEJCHANDRA S.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF014785; AAB67823.1; -

DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 83 83
 SQ SEQUENCE 83 AA; 9731 MW; 24B8D666 CRC32;

Query Match 76.0%; Score 19; DB 7; Length 83;
 Best Local Similarity 80.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 || ||
 Db 68 RESLR 72

RESULT 13
 ID O19541 PRELIMINARY; PRT; 83 AA.
 AC O19541;
 DT 01-JAN-1998 (TREMELrel. 05, Created)
 DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)
 DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
 RA RUNGRONG E., BEJCHANDRA S.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF014787; AAB67825.1; -
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 83 83
 SQ SEQUENCE 83 AA; 9731 MW; 24B8D666 CRC32;

Query Match 76.0%; Score 19; DB 7; Length 83;
 Best Local Similarity 80.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 || ||
 Db 68 RESLR 72

RESULT 14
 ID O19543 PRELIMINARY; PRT; 83 AA.
 AC O19543;
 DT 01-JAN-1998 (TREMELrel. 05, Created)
 DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)
 DE MHC CLASS I ANTIGEN HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
 RA RUNGRONG E., BEJCHANDRA S.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF014789; AAB67827.1; -
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 83 83
 SQ SEQUENCE 83 AA; 9731 MW; 24B8D666 CRC32;

Query Match 76.0%; Score 19; DB 7; Length 83;
 Best Local Similarity 80.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 ||||
 Db 68 RESLR 72

RESULT 15

O95956
 ID O95956 PRELIMINARY; PRT; 89 AA.
 AC O95956;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE HUMAN LEUCOCYTE ANTIGEN B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA BIRD T., DUNN P.P., BUNCE M.;
 RT "Three new HLA class I alleles in an individual from Kenya."
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ133471; CAB37940.1; -
 FT NON-TER 1
 FT NON-TER 89
 SQ SEQUENCE 89 AA; 10568 MW; 531E6106 CRC32;

Query Match 76.0%; Score 19; DB 4; Length 89;
 Best Local Similarity 80.0%; Pred. No. 3.7e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REXLR 5
 ||||
 Db 74 RESLR 78

Search completed: February 8, 2000, 13:17:28
 Job time: 32477 sec

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-3 to: GenEmbl.* out_format : pfs

Date: Feb 8, 2000 4:37 PM

About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=Cgml_1/USPTO.spool/US08653294/runat_04022000.160701_15779/app_query.fasta.1
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000
-GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=200 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294
-NCPU=6 -ICPU=3 -NO_XLPUX -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-3

Query length: 10

Database: GenEmbl.*

Database sequences: 821193

Database length: -1518192014

Search time (sec): 11370.480000

score_list:

| Sequence | Strd Orig | Zscore | Bscore | Len | Documentation |
|-------------------|-----------|--------|--------|---------|--------------------------------------|
| gb_pat:A40090 | + | 19.00 | 122.04 | 89.28 | 15 A40090 Sequence 51 from Patent W |
| gb_pat:A40091 | + | 19.00 | 121.53 | 95.36 | 16 A40091 Sequence 52 from Patent W |
| gb_pat:A40064 | + | 19.00 | 120.59 | 107.52 | 18 A40064 Sequence 25 from Patent W |
| gb_pat:R02H06 | - | 19.00 | 120.16 | 113.62 | 19 D00274 Homo sapiens gene for tyr |
| gb_pat:AR034425 | - | 19.00 | 118.31 | 144.17 | 24 AR034425 Sequence 5 from patent |
| gb_pat:AR017933 | + | 19.00 | 116.53 | 181.00 | 30 AR017933 Sequence 27 from patent |
| gb_pat:AR021017 | + | 19.00 | 114.05 | 248.88 | 41 AR021017 Sequence 50 from patent |
| gb_pat:AR043432 | + | 19.00 | 114.05 | 248.88 | 41 AR043432 Sequence 50 from patent |
| gb_pat:AR062347 | + | 19.00 | 114.05 | 248.88 | 41 AR062347 Sequence 50 from patent |
| gb_pat:E05342 | + | 19.00 | 113.31 | 273.65 | 45 E05342 Listeria monocytogenes sp |
| gb_pat:E06878 | + | 19.00 | 112.96 | 286.06 | 47 E06878 Synthetic DNA related to |
| gb_pat:S73031 | + | 19.00 | 112.80 | 292.26 | 48 S73031 TCR V alpha 8 (V alpha 8H |
| gb_pat:EL3614 | + | 19.00 | 110.89 | 373.16 | 61 EL3614 Polyrionucleotide with x |
| gb_pat:EL3623 | + | 19.00 | 110.64 | 385.64 | 63 EL3623 Polyrionucleotide with x |
| gb_pat:S401752 | + | 19.00 | 108.35 | 517.08 | 84 S40175 pancreatic cholesterol es |
| gb_pat:RSPES004 | + | 19.00 | 107.89 | 548.47 | 89 AL033574 H.sapiens flow-sorted c |
| gb_pat:I65704 | + | 19.00 | 107.29 | 592.49 | 96 I65704 Sequence 64 from patent |
| gb_pat:I67936 | + | 19.00 | 107.29 | 592.49 | 96 I67936 Sequence 64 from patent |
| gb_pat:I90157 | + | 19.00 | 107.29 | 592.49 | 96 I90157 Sequence 64 from patent |
| gb_pat:HUM2NFH | + | 19.00 | 106.96 | 617.67 | 100 M88364 Homo sapiens DNA-binding |
| gb_pat:HUM5SRRNA | + | 19.00 | 106.35 | 668.08 | 108 M35170 G.obcuriglobus 5S ribos |
| gb_pat:DMAANY | + | 19.00 | 106.06 | 693.32 | 112 X73658 D.melanogaster alpha-am |
| gb_pat:HS304F9R | + | 19.00 | 105.51 | 743.84 | 120 I79933 H.sapiens chromosome 22 |
| gb_pat:MUSMTOCB | + | 19.00 | 105.32 | 762.80 | 123 M8128 Mouse cardiac myosin hea |
| gb_pat:HS16568F | + | 19.00 | 105.19 | 775.45 | 125 I259607 H.sapiens CpG island DNA |
| gb_pat:HUMSG13 | + | 19.00 | 104.88 | 807.09 | 130 I05418 Human cell surface glyco |
| gb_pat:S72924 | + | 19.00 | 104.46 | 851.41 | 137 S72924 Ig VDJ gamma-immunoglob |
| gb_pat:PHIP30149M | + | 19.00 | 104.29 | 870.43 | 140 L34274 Plasmid Ophi membrane pr |
| gb_pat:S67638 | + | 19.00 | 104.23 | 876.77 | 141 S67638 HLA-B (HLA-B39.1)-HLA cl |
| gb_pat:HUMAI23 | + | 19.00 | 104.06 | 895.79 | 144 D28364 Human mRNA for annexin |
| gb_pat:R01472 | + | 19.00 | 103.74 | 933.86 | 150 E01472 DNA encoding gamma-gluta |
| gb_pat:AF136483 | + | 19.00 | 103.74 | 933.86 | 150 AF136483 Scare cereale clone S |
| gb_pat:G32681 | + | 19.00 | 103.58 | 952.90 | 153 G32681 A009N15 Human Homo sapie |
| gb_pat:HSSRC7 | + | 19.00 | 103.43 | 971.95 | 156 G03995 Human c-src gene exon 7 |
| gb_pat:AR042423 | + | 19.00 | 103.28 | 991.01 | 159 AR042423 Sequence 15 from patent |
| gb_pat:AR054638 | + | 19.00 | 103.28 | 991.01 | 159 AR054638 Sequence 15 from patent |
| gb_pat:HS30D1F | - | 19.00 | 103.23 | 997.37 | 160 I260668 H.sapiens CpG island DNA |
| gb_pat:EVM168NA7 | - | 19.00 | 102.93 | 1.0e+03 | 166 D31674 Enterococcus avium 16S r |
| gb_pat:HUMSWX1283 | - | 19.00 | 102.93 | 1.0e+03 | 166 L41954 Human chromosome X STS |
| gb_pat:HSK0X25 | + | 19.00 | 102.84 | 1.0e+03 | 168 X52356 Human Kox25 mRNA for zin |
| gb_pat:HS54B2R | + | 19.00 | 102.79 | 1.1e+03 | 169 Z55640 H.sapiens CpG island DNA |
| gb_ov:AB031373 | - | 19.00 | 102.74 | 1.1e+03 | 170 AB031373 Locustella pyretri DNA |

gb_pat:196192 + 19.00 102.74 1.1e+03 170 196192 Sequence 29 from pate
gb_pat:HS50D4R + 19.00 102.74 1.1e+03 170 255580 H.sapiens CpG island
gb_pat:HUMSRC06 + 19.00 102.74 1.1e+03 170 K03213 Human c-src-1 proto-o

seq_name: gb_pat:A40090

seq_documentation_block:

LOCUS A40090 15 bp DNA PAT 05-MAR-1997

DEFINITION Sequence 51 from Patent WO9421818.

ACCESSION A40090

VERSION A40090.1 GI:2296255

KEYWORDS

SOURCE

ORGANISM

unidentified.

unclassified.

REFERENCE

1 (bases 1 to 15)

Audrien,M., Dupont,E., Rossau,R. and De.C.I.

PROCESS FOR TYPING HLA-B USING SPECIFIC PRIMERS AND PROBES SETS

TITLE Patent: WO 9421818-A 51 29-SEP-1994;

JOURNAL INNOGENETICS NV (BE)

COMMENT Other publication CA 2158578 940929

Other publication AU 6258594 941011.

FEATURES

source

1..15

/organism="unidentified"

/db_xref="taxon:32644"

3 a 4 c 7 g 1 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 19.00 Length: 5

Ratio: 3.800 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x A40090 ..

Align seg 1/1 to: A40090 from: 1 to: 15

1 ArgGlu***LeuArg 5

|||||

1 CGAGAGAGCGCTGGG 15

seq_name: gb_pat:A40091

seq_documentation_block:

LOCUS A40091 16 bp DNA PAT 05-MAR-1997

DEFINITION Sequence 52 from Patent WO9421818.

ACCESSION A40091

VERSION A40091.1 GI:2296256

KEYWORDS

SOURCE

unidentified.

unclassified.

REFERENCE

1 (bases 1 to 16)

Audrien,M., Dupont,E., Rossau,R. and De.C.I.

PROCESS FOR TYPING HLA-B USING SPECIFIC PRIMERS AND PROBES SETS

TITLE Patent: WO 9421818-A 52 29-SEP-1994;

JOURNAL INNOGENETICS NV (BE)

COMMENT Other publication CA 2158578 940929

Other publication AU 6258594 941011.

FEATURES

source

1..16

/organism="unidentified"

/db_xref="taxon:32644"

4 a 4 c 7 g 1 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 19.00 Length: 5

Ratio: 3.800 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x A40091

Align seg 1/1 to: A40091 from: 1 to: 16

1 ArgGlu***LeuArg 5

|||||

1 CGAGAGAGCTGCGG 15

seq_name: gb_pat:A40064

seq_documentation_block:

LOCUS A40064 18 bp DNA PAT 05-MAR-1997

DEFINITION Sequence 25 from Patent WO9421818.

ACCESSION A40064

VERSION A40064.1 GI:2296229

KEYWORDS

SOURCE

ORGANISM

unidentified.

unclassified.

REFERENCE

AUTHORS

Andrien,M., Dupont,E., Rossau,R. and De,C.I.

TITLE

PROCESS FOR TYPING HLA-B USING SPECIFIC PRIMERS AND PROBES SETS

JOURNAL

Patent: WO 9421818-A 25 29-SEP-1994;

INNOGENETICS NV (BE)

COMMENT

Other publication CA 2158578 940929

Other publication AU 6258594 941011.

FEATURES

Location/Qualifiers

1..18

source

/organism="unidentified"

/db_xref="taxon:32644"

5 a 5 c 7 g 1 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 19.00

Ratio: 3.800

Length: 5

Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x A40064

Align seg 1/1 to: A40064 from: 1 to: 18

1 ArgGlu***LeuArg 5

|||||

3 CGAGAGAGCTGCGG 17

seq_name: gb_prl:HUMTH06

seq_documentation_block:

LOCUS HUMTH06 19 bp DNA PRI 11-MAR-1998

DEFINITION Homo sapiens gene for tyrosine hydroxylase, exon 6, partial

sequence.

ACCESSION D00274

VERSION D00274.1 GI:220104

KEYWORDS

TH: tyrosine hydroxylase.

SEGMENT 6 of 24

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

AUTHORS

Kobayashi,K., Kaneda,N., Ichinose,H., Kishi,F., Nakazawa,A.,

Kurosawa,Y., Fujita,K. and Nagatsu,T.

TITLE

Structure of the human tyrosine hydroxylase gene: alternative

splicing from a single gene accounts for generation of four mRNA

types

JOURNAL

J. Biochem. 103 (6), 907-912 (1988)

MEDLINE

89008200

REFERENCE

2 (sites)
O'Malley,K.L., Anhalt,M.J., Martin,B.M., Kelsoe,J.R., Winfield,S.L.
and Ginns,E.I.

TITLE

Isolation and characterization of the human tyrosine hydroxylase
gene: identification of 5' alternative splice sites responsible for
multiple mRNAs

JOURNAL

Biochemistry 26, 6910-6914 (1987)

COMMENT

In [1], they determined the nucleotide sequences of all exons and
their surrounding regions of human TH gene, and the exon/intron
boundaries are shown. The boundaries were determined by comparing
the genomic DNA sequence with the cDNA sequence. The human TH gene
is split into 14 exons. In [1], they concluded that the four types
of human TH mRNA are produced through alternative splicing from a
single gene.

FEATURES

Location/Qualifiers

1..19

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="gHTH-E20"

/tissue_type="placenta"

/note="157 bp after segment 5"

<1..10

/number=5

11..>19

/note="AA 224 at 11"

/number=6

/product="tyrosine hydroxylase"

1 a 8 c 5 g 5 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 19.00

Ratio: 3.800

Length: 5

Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x HUMTH06/rev ..

Align seg 1/1 to reverse of: HUMTH06 from: 1 to: 19

1 ArgGlu***LeuArg 5

|||||

19 CGAGAAGCCTGAGG 5

seq_name: gb_pat:AR034425

seq_documentation_block:

LOCUS AR034425 24 bp DNA PAT 29-SEP-1999

DEFINITION Sequence 5 from patent US 5869438.

ACCESSION AR034425

VERSION AR034425.1 GI:5950030

KEYWORDS

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 24)

Svendsen,A., Patkar,S.Anant, Gormsen,E., Okkels,J.Sigurd and

Thellersen,M.

Lipase variants

JOURNAL

Patent: US 5869438-A 5 09-FEB-1999;

FEATURES

Location/Qualifiers

1..24

source

/organism="unknown"

4 a 4 c 9 g 7 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 19.00

Ratio: 3.800

Length: 5

Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x AR034425/rev ..

Align seg 1/1 to reverse of: AR034425 from: 1 to: 24

1 ArgGlu***LeuArg 5
|||||

17 CGAGAATCCCTGCGC 3

seq_name: gb_pat:AR017933

seq_documentation_block:

LOCUS AR017933 30 bp DNA 05-DEC-1998 PAT

DEFINITION Sequence 27 from patent US 5780272.

ACCESSION AR017933

VERSION AR017933.1 GI:3973536

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 30)

AUTHORS Jarrell,K.A.

TITLE Intron-mediated recombinant techniques and reagents

JOURNAL Patent: US 5780272-A 27 14-JUL-1998;

FEATURES Location/Qualifiers.

source 1..30

BASE COUNT 8 a 8 c 8 g 6 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x AR017933 ..

Align seg 1/1 to: AR017933 from: 1 to: 30

1 ArgGlu***LeuArg 5
|||||

15 ACAGAAGCACTGCGC 29

seq_name: gb_pat:AR021017

seq_documentation_block:

LOCUS AR021017 41 bp DNA 05-DEC-1998 PAT

DEFINITION Sequence 50 from patent US 5789245.

ACCESSION AR021017

VERSION AR021017.1 GI:3975632

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 41)

AUTHORS Dubensky,T.W. Jr., Polo,J.M., Ibanez,C.E., Chang,S.M.W., Jolly,D.J.

TITLE Alphavirus structural protein expression cassettes

JOURNAL Patent: US 5789245-A 50 04-AUG-1998;

FEATURES Location/Qualifiers

source 1..41

BASE COUNT 17 a 6 c 5 g 13 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x AR021017 ..

Align seg 1/1 to: AR021017 from: 1 to: 41

1 ArgGlu***LeuArg 5
|||||

10 CGAGAAGCTCTAAGG 24

seq_name: gb_pat:AR043432

seq_documentation_block:

LOCUS AR043432 41 bp DNA 29-SEP-1999 PAT

DEFINITION Sequence 50 from patent US 5814482.

ACCESSION AR043432

VERSION AR043432.1 GI:5964440

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 41)

AUTHORS Dubensky,T.W. Jr., Polo,J.M., Jolly,D.J. and Driver,D.A.

TITLE Eukaryotic layered vector initiation systems

JOURNAL Patent: US 5814482-A 50 29-SEP-1998;

FEATURES Location/Qualifiers

source 1..41

BASE COUNT 17 a 6 c 5 g 13 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x AR043432 ..

Align seg 1/1 to: AR043432 from: 1 to: 41

1 ArgGlu***LeuArg 5
|||||

10 CGAGAAGCTCTAAGG 24

seq_name: gb_pat:AR062347

seq_documentation_block:

LOCUS AR062347 41 bp DNA 29-SEP-1999 PAT

DEFINITION Sequence 50 from patent US 5843723.

ACCESSION AR062347

VERSION AR062347.1 GI:5990038

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 41)

AUTHORS Dubensky,T.W. Jr., Polo,J.M., Ibanez,C.E., Chang,S.M.W.,

Jolly,D.J., Driver,D.A. and Belli,B.A.

TITLE Alphavirus vector constructs

JOURNAL Patent: US 5843723-A 50 01-DEC-1998;

FEATURES Location/Qualifiers

source 1..41

BASE COUNT 17 a 6 c 5 g 13 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:
  US-08-653-294-3 x AR062347 ..
  Align seg 1/1 to: AR062347 from: 1 to: 41
    1 ArgGlu***LeuArg 5
    10 CGAGAAGCTCTAAGG 24
  seq_name: gb_pat:E05342
  seq_documentation_block:
    LOCUS E05342 45 bp RNA PAT 29-SEP-1997
    DEFINITION Listeria monocytogenes-specific probe.
    ACCESSION E05342
    VERSION E05342.1 GI:2173531
    KEYWORDS Jp 1993219997-A/4.
    SOURCE unidentified.
    ORGANISM unclassified.
    REFERENCES 1 (bases 1 to 45)
    AUTHORS Hayashi,T., Yoshida,M. and Terada,K.
    TITLE PROBE FOR DETECTING LISTERIA AND METHOD FOR DETECTION
    JOURNAL Patent: JP 1993219997-A 4 31-AUG-1993;
    TOAGOSEI CHEM IND CO LTD
    COMMENT PN JP 1993219997-A/4
    PD 31-AUG-1993
    PF 04-FEB-1992 JP 1992047959
    PI HAYASHI TAKAKO, YOSHIDA MASAO, TERADA KENJI
    PC C12Q1/68;
    CC CC strandedness: Single;
    CC topology: Linear;
    CC anti-sense: No;
    CC Location/Qualifiers
      1..45
      /organism="unidentified"
      /db_xref="taxon:32644" 12 t
  BASE COUNT 8 a 15 c 10 g 12 t
  ORIGIN
    1..45
    /organism="unidentified"
    /db_xref="taxon:32644" 12 t
  alignment_scores:
    Quality: 19.00 Length: 5
    Ratio: 3.800 Gaps: 0
    Percent Similarity: 100.000 Percent Identity: 100.000
  alignment_block:
    US-08-653-294-3 x AR062347 ..
    Align seg 1/1 to: AR062347 from: 1 to: 41
      1 ArgGlu***LeuArg 5
      10 CGAGAAGCTCTAAGG 24
    seq_name: gb_pr2:S73031
    seq_documentation_block:
      LOCUS S73031 48 bp DNA PRI 26-JAN-1995
      DEFINITION TCR V alpha 8 (V alpha 8AP50)-T-cell receptor alpha chain variable
      region [human, peripheral blood lymphocytes, bullous pemphigoid
      patient, Genomic, 48 nt].
      ACCESSION S73031
      VERSION S73031
      KEYWORDS S73031.1 GI:639537
      SOURCE human bullous pemphigoid patient peripheral blood lymphocytes.
      ORGANISM Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
      Eutheria; Primates; Catarrhini; Hominidae; Homo.
      REFERENCE 1 (bases 1 to 48)
      AUTHORS Michalaki,H., Roman-Roman,S., Nicolas,J.F., Mackensen,A.,
      Thivolet,J., Triebel,F., Hercend,T. and Ferradini,L.
      TITLE In-situ preferential usage of V alpha 8 T-cell receptor gene
      segments in a patient with bullous pemphigoid
      JOURNAL J. Autoimmun. 6 (6), 827-839 (1993)
      MEDLINE 94206429
      REMARK GenBank staff at the National Library of Medicine created this
      entry [NCBI gbsq 154951] from the original journal article.
      This sequence comes from Fig. 3C.
    FEATURES
      source
      1..48
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      gene
      1..48
      /partial
      /gene="TCR Vαgr:8"
      /note="T-cell receptor alpha chain variable region"
      /allele="Vαgr:8AP50"
      CDS
      1..48
      /partial
      /gene="TCR Vαgr:8"
      /note="This sequence comes from Fig. 3C."
      /codon_start=1
      /product="T-cell receptor alpha chain variable region"
      /protein_id="AAB30538.1"
      /db_xref="GI:639538"
      /translation="CAENLYGGGSEKLFG"
  
```

BASE COUNT 12 a 8 c 16 g 12 t
ORIGIN

alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x S73031 ..

Align seg 1/1 to: S73031 from: 1 to: 48

1 ArgGlu**LeuArg 5

|||||
6 AGAGAACTCTACGG 20

seq_name: gb_pat:EI3614

seq_documentation_block: 61 bp RNA PAT 27-APR-1998
LOCUS EI3614
DEFINITION Polyribonucleotide with ribozyme activity which has a tRNA
anti-codon stem loop.

ACCESSION EI3614

VERSION EI3614.1 GI:5708660

KEYWORDS JP 1997220094-A/2.

SOURCE unidentified.

ORGANISM unclassified.

REFERENCE 1 (bases 1 to 61)

AUTHORS Koizumi,M., Ozawa,Y. and Nishigaki,T.

TITLE RIBOZYME HAVING TRNA ANTICODON STEM LOOP

JOURNAL Patent: JP 1997220094-A 26-AUG-1997;

COMMENT SANKYO CO LTD

OS None

OC Artificial sequences.

PN JP 1997220094-A/2

PD 26-AUG-1997

PF 12-DEC-1996 JP 1996331843

PR 13-DEC-1995 JP 95P 324778

PI KOIZUMI MAKOTO, OZAWA YUJI, NISHIGAKI TAKASHI PC

Cl2N15/09,C07H21/02,C07H21/04,C12N5/10,C12N9/00,(C12N15/09, PC

Cl2R1:92);

CC strandedness: Single;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

FH Key Location/Qualifiers

FT source 1..61

FT stem_loop 16..60. /organism='Artificial sequences' FT

FEATURES

source 1..61 Location/Qualifiers

/organism='unidentified'

/db_xref="taxon:32644"

BASE COUNT 24 a 13 c 11 g 13 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x EI3614 ..

Align seg 1/1 to: EI3614 from: 1 to: 61

1 ArgGlu**LeuArg 5

|||||

19 AGAGAAACACTCAGA 33

seq_name: gb_pat:EI3623

seq_documentation_block:

LOCUS EI3623 63 bp RNA PAT 27-APR-1998

DEFINITION Polyribonucleotide with ribozyme activity which has a tRNA

anti-codon stem loop.

ACCESSION EI3623

VERSION EI3623.1 GI:5708669

KEYWORDS JP 1997220094-A/11.

SOURCE unidentified.

ORGANISM unclassified.

REFERENCE 1 (bases 1 to 63)

AUTHORS Koizumi,M., Ozawa,Y. and Nishigaki,T.

TITLE RIBOZYME HAVING TRNA ANTICODON STEM LOOP

JOURNAL Patent: JP 1997220094-A 26-AUG-1997;

COMMENT SANKYO CO LTD

OS None

OC Artificial sequences.

PN JP 1997220094-A/11

PD 26-AUG-1997

PF 12-DEC-1996 JP 1996331843

PR 13-DEC-1995 JP 95P 324778

PI KOIZUMI MAKOTO, OZAWA YUJI, NISHIGAKI TAKASHI PC

Cl2N15/09,C07H21/02,C07H21/04,C12N5/10,C12N9/00,(C12N15/09, PC

Cl2R1:92);

CC strandedness: Single;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

FH Key Location/Qualifiers

FT source 1..63

FT stem_loop 18..62. /organism='Artificial sequences' FT

FEATURES

source 1..63 Location/Qualifiers

/organism='unidentified'

/db_xref="taxon:32644"

BASE COUNT 21 a 13 c 12 g 17 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x EI3623 ..

Align seg 1/1 to: EI3623 from: 1 to: 63

1 ArgGlu**LeuArg 5

|||||

21 AGAGAAACACTCAGA 35

seq_name: gb_pr2:S40177S2

seq_documentation_block:

LOCUS S40177S2 84 bp DNA PRI 08-MAY-1993

DEFINITION pancreatic cholesterol esterase, [5' region, exon 1, exon 10]

[human, pWE 15, PTCF, Genomic, 84 nt, segment 2 of 2].

ACCESSION S40178

VERSION S40178.1 GI:251572

KEYWORDS

SEGMENT 2 of 2

SOURCE human PTCF pWE 15.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 84)
 AUTHORS Kumar,B.V., Aleman-Gomez,J.A., Colwell,N., Lopez-Candales,A.,
 Bosner,M.S., Spillburg,C.A., Lowe,M. and Lange,L.G.
 TITLE Structure of the human pancreatic cholesterol esterase gene
 JOURNAL Biochemistry 31 (26), 6077-6081 (1992)
 MEDLINE 92329425
 REMARK GenBank staff at the National Library of Medicine created this
 entry [NCBI gibbsg 109184] from the original journal article.
 This sequence comes from Table II.

FEATURES
 source
 1..84 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 gene
 1..84 /partial
 /gene="pancreatic cholesterol esterase, CEase"
 CDS
 1..84 /partial
 /gene="pancreatic cholesterol esterase, CEase"
 /note="precursor. This sequence comes from Table II;
 CEase"
 /codon_start=1
 /product="pancreatic cholesterol esterase"
 /protein_id="AAH2537.1"
 /db_xref="GI:251573"
 /translation="HADDIQYVFGKPFATPTGYRPODRTVSK"
 BASE COUNT 20 a 29 c 20 g 15 t
 ORIGIN

alignment_scores:
 Quality: 19.00 Length: 5
 Ratio: 3.800 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x S40177S2 ..

Align seg 1/1 to: S40177S2 from: 1 to: 84

1 ArgGlu***LeuArg 5
 |||||
 27 CGGGAAGCCCTTCGC 41

OM of: US-08-653-294-3 to: N_Geneseq_36.* out_format : pfs
 Date: Feb 8, 2000 1:27 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
 -MODEL=frameat_p2n.model -DEV=xlp
 -Q=/cgnl1/USPTO_spool/US08653294/runat_04022000_160701_15807/app_query.fasta.1
 -DB=N_Geneseq_36 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000
 -GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -GAPOP=6.000 -GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT
 -THREADS=1

Search information block:
 Query: US-08-653-294-3
 Query length: 10
 Database: N_Geneseq_36.*
 Database sequences: 311585
 Database length: 125096042
 Search time (sec): 590.520000

score_list:

| Sequence | Strd | Orig | ZScore | EScore | Len | Documentation |
|----------------------|------|-------|--------|---------|-----|------------------------------------|
| N_Geneseq_36:Q72680 | + | 19.00 | 119.14 | 48.97 | 15 | Probe S19 for distinguishing HLA-B |
| N_Geneseq_36:Q72681 | + | 19.00 | 118.62 | 52.35 | 16 | Probe S19(2) for distinguishing |
| N_Geneseq_36:Q72654 | + | 19.00 | 117.67 | 59.14 | 18 | Probe 19, anneals to bases 220-2 |
| N_Geneseq_36:Q87369 | + | 19.00 | 113.55 | 100.37 | 30 | TPA PCR primer. New intron-media |
| N_Geneseq_36:V67317 | + | 19.00 | 113.55 | 100.37 | 30 | PCR primer used to amplify human |
| N_Geneseq_36:V68669 | + | 19.00 | 113.55 | 100.37 | 30 | PCR primer used to amplify human |
| N_Geneseq_36:V03294 | - | 19.00 | 112.54 | 114.25 | 34 | Isolate OC9a phosphatase (27A3A) |
| N_Geneseq_36:V82702 | - | 19.00 | 112.54 | 114.25 | 34 | Antisense PCR primer C2y used am |
| N_Geneseq_36:V30827 | + | 19.00 | 111.03 | 138.69 | 41 | XhoI insertion primer XSVS04647F |
| N_Geneseq_36:V42404 | + | 19.00 | 111.03 | 138.69 | 41 | Forward PCR primer XSVS04647F |
| N_Geneseq_36:V60164 | + | 19.00 | 111.03 | 138.69 | 41 | Forward primer XSVS04647F used t |
| N_Geneseq_36:V07024 | - | 19.00 | 111.03 | 138.69 | 41 | Forward primer XSVS04647F used t |
| N_Geneseq_36:Q48012 | - | 19.00 | 110.28 | 152.72 | 45 | Listeria monocytogenes DNA detec |
| N_Geneseq_36:Q58576 | - | 19.00 | 110.28 | 152.72 | 45 | Listeria monocytogenes RNA detec |
| N_Geneseq_36:Q63760 | + | 19.00 | 109.82 | 159.75 | 47 | Mink growth hormone/A. oryzae ne |
| N_Geneseq_36:V92326 | + | 19.00 | 107.92 | 209.26 | 61 | Ribozyme (2) to specifically cle |
| N_Geneseq_36:V92335 | + | 19.00 | 107.56 | 216.36 | 63 | Ribozyme (11) to specifically cle |
| N_Geneseq_36:Q56719 | + | 19.00 | 104.50 | 320.21 | 92 | Sequence of chick beta-actin pro |
| N_Geneseq_36:V65277 | + | 19.00 | 104.16 | 334.64 | 96 | Platelet derived growth factor |
| N_Geneseq_36:V13651 | - | 19.00 | 101.23 | 487.25 | 138 | Hepatitis G virus nucleic acid |
| N_Geneseq_36:V53893 | + | 19.00 | 100.09 | 564.21 | 159 | Receptor protein tyrosine kinase |
| N_Geneseq_36:V65315 | + | 19.00 | 100.09 | 564.21 | 159 | Receptor protein tyrosine kinase |
| N_Geneseq_36:V20451 | + | 19.00 | 99.55 | 604.67 | 170 | Human c-src oncogene. Anticand |
| N_Geneseq_36:Q77371 | + | 19.00 | 98.91 | 656.30 | 182 | Human genome fragment (Preferre |
| N_Geneseq_36:V40976 | + | 19.00 | 98.56 | 685.86 | 194 | Human secreted protein 5' EST |
| N_Geneseq_36:Q53267 | + | 19.00 | 98.40 | 700.66 | 196 | Human brain Expressed Sequence |
| N_Geneseq_36:V161620 | + | 19.00 | 97.46 | 789.17 | 220 | Lactobacillus delbrueckii ssp. |
| N_Geneseq_36:V161621 | + | 19.00 | 97.39 | 797.11 | 222 | Lactobacillus delbrueckii ssp. |
| N_Geneseq_36:V161622 | + | 19.00 | 97.39 | 797.11 | 222 | L. delbrueckii ssp. bulgaricus |
| N_Geneseq_36:V39550 | + | 19.00 | 97.39 | 797.11 | 222 | Human secreted protein 5' EST |
| N_Geneseq_36:V21403 | + | 19.00 | 97.18 | 819.42 | 228 | Human gene signature HUMG02772 |
| N_Geneseq_36:V51592 | + | 19.00 | 96.34 | 912.62 | 253 | Human secreted protein 5' EST |
| N_Geneseq_36:Q61237 | + | 19.00 | 95.96 | 957.47 | 265 | Human brain Expressed Sequence |
| N_Geneseq_36:V86668 | - | 19.00 | 95.93 | 961.21 | 266 | EST clone Bg248. New polynucleo |
| N_Geneseq_36:V87265 | - | 19.00 | 95.58 | 1.0e+03 | 278 | Arabidopsis thaliana antifungal |
| N_Geneseq_36:V83612 | + | 19.00 | 95.35 | 1.0e+03 | 286 | Nucleotide sequence of clone GM |
| N_Geneseq_36:V81378 | + | 19.00 | 95.35 | 1.0e+03 | 286 | MSRV-1 virus clone GM3. Peptide |
| N_Geneseq_36:V89988 | - | 19.00 | 95.12 | 1.1e+03 | 294 | EST clone CW675. New polynucleo |
| N_Geneseq_36:V20162 | - | 19.00 | 95.07 | 1.1e+03 | 296 | Probe (63) for microbial genes |
| N_Geneseq_36:Q79120 | - | 19.00 | 94.93 | 1.1e+03 | 301 | Osteocalcin gene 5' region. New |
| N_Geneseq_36:V88322 | + | 19.00 | 94.83 | 1.1e+03 | 305 | EST clone FX536. New polynucleo |

| | | | | | | |
|---------------------|---|-------|-------|---------|-----|------------------------------|
| N_Geneseq_36:Q53891 | - | 19.00 | 94.49 | 1.2e+03 | 318 | Expressed Sequence Tag human |
| N_Geneseq_36:Q59303 | - | 19.00 | 94.49 | 1.2e+03 | 318 | Human brain Expressed Sequen |
| N_Geneseq_36:V35857 | - | 19.00 | 94.36 | 1.2e+03 | 323 | cDNA encoding protein homolo |
| N_Geneseq_36:V95954 | - | 19.00 | 94.36 | 1.2e+03 | 323 | Arabidopsis lysine ketogluta |

seq_name: N_Geneseq_36:Q72680

seq documentation_block:

| | | |
|----|---|-----------------------|
| ID | Q72680 | standard; DNA; 15 BP. |
| AC | Q72680; | |
| DE | 22-MAY-1995 (first entry) | |
| DE | Probe S19 for distinguishing HLA-B alleles. | |
| KW | Polymerase chain reaction; PCR; primer; amplif; detection: HLA-B; | |
| KW | human leukocyte antigen; B25P; allele; exon 2; B54(22); B52(5); B7801; | |
| KW | B62(15); B75(15); B71(70); B46; B79; B53; B5102; B5103; B58(17); | |
| KW | compatibility; donor; recipient; organ transplant; success-rate; | |
| KW | bone-marrow transplant; disease susceptibility study; probe; | |
| KW | forensic investigation; ss. | |
| OS | Synthetic. | |
| PN | WO9421818-A. | |
| PD | 29-SEP-1994. | |
| PF | 07-MAR-1994; E00654. | |
| PR | 18-MAR-1993; EP-400700. | |
| PA | (INNO-) INNOGENETICS NV SA. | |
| PI | Andrien M, De Canck I, Dupont E, Rossau R; | |
| DR | WPI: 94-317037/39. | |
| PT | DNA typing using primers and probes enabling discrimination of | |
| PT | HLA-B alleles - esp. where difficult to discriminate by | |
| PT | serological means. | |
| PS | Claim 17; Page 50; 66pp; English. | |
| CC | The sequences given in Q72636-55 are probes which were used to | |
| CC | discriminate HLA-B alleles which are characterised by having the | |
| CC | sequence 5'-GCCA-3' at position 30-33 of exon 2 of the HLA-B allele | |
| CC | variants of these sequences which may also be used are given in Q72656- | |
| CC | 81. These probes are used to identify the amplification products of the | |
| CC | primers given in Q72630-35. These primers may be used to distinguish | |
| CC | between HLA-B types which are serologically difficult to discriminate, | |
| CC | eg. B54(22), B52(5), B7801, B62(15), B75(15), B71(70), B46, B79, B53, | |
| CC | B5102, B5103 and B58(17). Using this method, HLA-B compatibility | |
| CC | between donors and recipients of organ transplants can be increased, | |
| CC | thereby having a beneficial impact on success-rate of organ and bone- | |
| CC | marrow transplants. HLA-B typing may be used to improve or facilitate | |
| CC | disease susceptibility studies and forensic investigations. | |
| SQ | Sequence 15 BP; 3 A; 4 C; 7 G; 1 T; | |

alignment_scores:

Quality: 19.00 Length: 5

Ratio: 3.800 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x Q72680 ..

Align seg 1/1 to: Q72680 from: 1 to: 15

1 ArgGlu***LeuArg 5

|||||

1 CGAGAGAGCTCGCG 15

seq_name: N_Geneseq_36:Q72681

seq documentation_block:

| | | |
|----|--|-----------------------|
| ID | Q72681 | standard; DNA; 16 BP. |
| AC | Q72681; | |
| DE | 22-MAY-1995 (first entry) | |
| DE | Probe S19(2) for distinguishing HLA-B alleles. | |
| KW | Polymerase chain reaction; PCR; primer; amplif; detection: HLA-B; | |
| KW | human leukocyte antigen; B25P; allele; exon 2; B54(22); B52(5); B7801; | |
| KW | B62(15); B75(15); B71(70); B46; B79; B53; B5102; B5103; B58(17); | |
| KW | compatibility; donor; recipient; organ transplant; success-rate; | |
| KW | bone-marrow transplant; disease susceptibility study; probe; | |
| KW | forensic investigation; ss. | |

```

OS Synthetic.
PN W09421818-A.
PD 29-SEP-1994.
PF 07-MAR-1994; E00654.
PR 18-MAR-1993; EP-400700.
PA (INNO-) INNOGENETICS NV SA.
PI Andrien M, De Canck I, Dupont E, Rossau R;
DR WPI; 94-317037/39.
PT DNA typing using primers and probes enabling discrimination of
PT HLA-B alleles - esp. where difficult to discriminate by
PT serological means.
PS Claim 17; Page 50; 66pp; English.
CC The sequences given in Q72636-55 are probes which were used to
CC discriminate HLA-B alleles which are characterised by having the
CC sequence 5'-GCCA-3' at position 30-33 of exon 2 of the HLA-B allele.
CC Variants of these sequences which may also be used are given in Q72656-
CC 81. These probes are used to identify the amplification products of the
CC primers given in Q72630-35. These primers may be used to distinguish
CC between HLA-B types which are serologically difficult to discriminate,
CC eg. B54(22), B52(5), B7801, B62(15), B75(15), B71(70), B46, B79, B53,
CC B5102, B5103 and B58(17). Using this method, HLA-B compatibility
CC between donors and recipients of organ transplants can be increased,
CC thereby having a beneficial impact on success-rate of organ and bone-
CC marrow transplants. HLA-B typing may be used to improve or facilitate
CC disease susceptibility studies and forensic investigations.
SQ Sequence 16 BP; 4 A; 4 C; 7 G; 1 T;

alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-3 x Q72681 ..
Align seg 1/1 to: Q72681 from: 1 to: 16

seq_name: N_Geneseq_36:Q72654
seq_documentation_block:
ID Q72654 standard; DNA; 18 BP.
AC Q72654;
DT 22-MAY-1995 (first entry)
DE Probe 19, anneals to bases 220-237 of exon 2 of HLA-B*4001.
KW Polymerase chain reaction; PCR; primer; amplification; detection; HLA-B;
KW human leukocyte antigen; B25p; allele; exon 2; B54(22); B52(5); B7801;
KW B62(15); B75(15); B71(70); B46; B79; B53; B5102; B5103; B58(17);
KW compatibility; donor; recipient; organ transplant; success-rate;
KW bone-marrow transplant; disease susceptibility study; probe;
KW forensic investigation; ss.
OS Synthetic.
PN W09421818-A.
PD 29-SEP-1994.
PF 07-MAR-1994; E00654.
PR 18-MAR-1993; EP-400700.
PA (INNO-) INNOGENETICS NV SA.
PI Andrien M, De Canck I, Dupont E, Rossau R;
DR WPI; 94-317037/39.
PT DNA typing using primers and probes enabling discrimination of
PT HLA-B alleles - esp. where difficult to discriminate by
PT serological means.
PS Claim 7; Page 42; 66pp; English.
CC The sequences given in Q72636-55 are probes which were used to
CC discriminate HLA-B alleles which are characterised by having the
CC sequence 5'-GCCA-3' at position 30-33 of exon 2 of the HLA-B allele.
CC Variants of these sequences which may also be used are given in Q72656-
CC 81. These probes are used to identify the amplification products of the
CC primers given in Q72630-35. These primers may be used to distinguish

```

```

CC between HLA-B types which are serologically difficult to discriminate,
CC eg. B54(22), B52(5), B7801, B62(15), B75(15), B71(70), B46, B79, B53,
CC B5102, B5103 and B58(17). Using this method, HLA-B compatibility
CC between donors and recipients of organ transplants can be increased,
CC thereby having a beneficial impact on success-rate of organ and bone-
CC marrow transplants. HLA-B typing may be used to improve or facilitate
CC disease susceptibility studies and forensic investigations.
SQ Sequence 18 BP; 5 A; 5 C; 7 G; 1 T;

alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-3 x Q72654 ..
Align seg 1/1 to: Q72654 from: 1 to: 18

seq_name: N_Geneseq_36:Q87369
seq_documentation_block:
ID Q87369 standard; DNA; 30 BP.
AC Q87369;
DT 19-SEP-1995 (first entry)
DE TPA PCR primer.
KW Intron; recombination; combinatorial gene; trans-splicing;
KW gene therapy; polymerase chain reaction; PCR; primer; amplification;
KW tissue plasminogen activator; tPA; plasmid TPA-KS+; thrombolytic;
KW ss.
OS Synthetic.
PN W09507351-A.
PD 16-MAR-1995.
PF 12-SEP-1994; U10146.
PR 10-SEP-1993; US-119512.
PA (HARD ) HARVARD COLLEGE.
PI Jarrell KA;
DR WPI; 95-123425/16.
PT New intron-mediated recombinant techniques - used for the
PT generation and selection of novel genes and gene prods. for use
PT in therapy
PS Example 4; Page 50; 87pp; English.
CC A cDNA clone of human tissue plasminogen activator (tPA) was
CC amplified by PCR using the primers given in Q87368-69. The
CC amplified tPA DNA (Q87370) was ligated into vector KS+ to
CC obtain plasmid TPA-KS+. The construct was used in combinatorial
CC methods involving RNA splicing-mediated shuffling of tPA domains
CC in plasmid pINVI (Q87347) to generate novel tPAs having
CC improved thrombolytic properties.
SQ Sequence 30 BP; 8 A; 8 C; 8 G; 6 T;

alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-3 x Q87369 ..
Align seg 1/1 to: Q87369 from: 1 to: 30

seq_name: N_Geneseq_36:V37317

```

```

seq_documentation_block:
ID V37317 standard; DNA; 30 BP.
AC V37317;
DT 10-SEP-1998 (first entry)
DE PCR primer used to amplify human tissue plasminogen activator gene.
KW Plasmid pINVI; reverse-splicing intron; group II intron;
KW exon binding site; domain V motif; branch site acceptor;
KW nucleophilic group; transesterification; phosphodiester bond;
KW autocatalytic Y-branched intron; reverse splicing reaction;
KW PCR primer; ss.
OS Synthetic.
OS Homo sapiens.
PN US5780272-A.
PD 14-JUL-1998.
PR 07-JUN-1995; 488015.
PR 10-SEP-1993; US-119512.
PA (HARD.) HARVARD COLLEGE.
PI Jarrell KA;
DR WPI; 98-413060/35.
PT Reverse splicing construct containing fragments of autocatalytic
PT introns - able to cleave and ligate discontinuous nucleic acid for
PT generating new genes and e.g. ribozymes, libraries of enzymes and
PT antibodies
PS Example 4; Column 42; 56pp; English.
CC PCR primers V37316-17 were used to amplify the human tissue plasminogen
CC activator gene. The PCR product was used to construct plasmid tPA-KS+,
CC which is used in the course of the invention. The specification
CC describes a purified reverse-splicing intron which comprises a segment
CC comprising a 5'-part of a group II intron, including an exon binding
CC site not naturally present in the intron and a second segment comprising
CC a 3'-part of a group II intron, including a domain V motif, a branch
CC site acceptor, and a nucleophilic group for transesterifying a
CC phosphodiester bond of an RNA. Together the two segments form an
CC autocatalytic Y-branched intron which catalyses integration of at least
CC the first segment into substrate RNA by a reverse splicing reaction
CC The reverse-splicing introns are used, by specific cleavage and ligation
CC of discontinuous nucleic acid, to generate new genes and gene products,
CC e.g. ribozymes (for use in gene therapy or as reagents in DNA
CC manipulation, e.g. replacements for restriction enzymes) or
CC immunologically active or signal-transducing proteins such as antibody
CC and enzyme libraries.
SQ Sequence 30 BP; 8 A; 8 C; 8 G; 6 T;

alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-3 x V37317
Align seg 1/1 to: V37317 from: 1 to: 30

seq_name: N_Geneseq_36:V37317
seq_documentation_block:
ID V37317 standard; DNA; 30 BP.
AC V37317;
DT 10-SEP-1998 (first entry)
DE PCR primer used to amplify human tissue plasminogen activator gene.
KW Plasmid pINVI; reverse-splicing intron; group II intron;
KW exon binding site; domain V motif; branch site acceptor;
KW nucleophilic group; transesterification; phosphodiester bond;
KW autocatalytic Y-branched intron; reverse splicing reaction;
KW PCR primer; ss.
OS Synthetic.
OS Homo sapiens.
PN US5780272-A.
PD 14-JUL-1998.
PR 07-JUN-1995; 488015.
PR 10-SEP-1993; US-119512.
PA (HARD.) HARVARD COLLEGE.
PI Jarrell KA;
DR WPI; 98-413060/35.
PT Reverse splicing construct containing fragments of autocatalytic
PT introns - able to cleave and ligate discontinuous nucleic acid for
PT generating new genes and e.g. ribozymes, libraries of enzymes and
PT antibodies
PS Example 4; Column 42; 56pp; English.
CC PCR primers V37316-17 were used to amplify the human tissue plasminogen
CC activator gene. The PCR product was used to construct plasmid tPA-KS+,
CC which is used in the course of the invention. The specification
CC describes a purified reverse-splicing intron which comprises a segment
CC comprising a 5'-part of a group II intron, including an exon binding
CC site not naturally present in the intron and a second segment comprising
CC a 3'-part of a group II intron, including a domain V motif, a branch
CC site acceptor, and a nucleophilic group for transesterifying a
CC phosphodiester bond of an RNA. Together the two segments form an
CC autocatalytic Y-branched intron which catalyses integration of at least
CC the first segment into substrate RNA by a reverse splicing reaction
CC The reverse-splicing introns are used, by specific cleavage and ligation
CC of discontinuous nucleic acid, to generate new genes and gene products,
CC e.g. ribozymes (for use in gene therapy or as reagents in DNA
CC manipulation, e.g. replacements for restriction enzymes) or
CC immunologically active or signal-transducing proteins such as antibody
CC and enzyme libraries.
SQ Sequence 30 BP; 8 A; 8 C; 8 G; 6 T;

alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-3 x V37317
Align seg 1/1 to: V37317 from: 1 to: 30

seq_name: N_Geneseq_36:V37317
seq_documentation_block:
ID V37317 standard; DNA; 30 BP.
AC V37317;
DT 10-SEP-1998 (first entry)
DE PCR primer used to amplify human tissue plasminogen activator gene.
KW Plasmid pINVI; reverse-splicing intron; group II intron;
KW exon binding site; domain V motif; branch site acceptor;
KW nucleophilic group; transesterification; phosphodiester bond;
KW autocatalytic Y-branched intron; reverse splicing reaction;
KW PCR primer; ss.
OS Synthetic.
OS Homo sapiens.
PN US5780272-A.
PD 14-JUL-1998.
PR 07-JUN-1995; 488015.
PR 10-SEP-1993; US-119512.
PA (HARD.) HARVARD COLLEGE.
PI Jarrell KA;
DR WPI; 98-413060/35.
PT Reverse splicing construct containing fragments of autocatalytic
PT introns - able to cleave and ligate discontinuous nucleic acid for
PT generating new genes and e.g. ribozymes, libraries of enzymes and
PT antibodies
PS Example 4; Column 42; 56pp; English.
CC PCR primers V37316-17 were used to amplify the human tissue plasminogen
CC activator gene. The PCR product was used to construct plasmid tPA-KS+,
CC which is used in the course of the invention. The specification
CC describes a purified reverse-splicing intron which comprises a segment
CC comprising a 5'-part of a group II intron, including an exon binding
CC site not naturally present in the intron and a second segment comprising
CC a 3'-part of a group II intron, including a domain V motif, a branch
CC site acceptor, and a nucleophilic group for transesterifying a
CC phosphodiester bond of an RNA. Together the two segments form an
CC autocatalytic Y-branched intron which catalyses integration of at least
CC the first segment into substrate RNA by a reverse splicing reaction
CC The reverse-splicing introns are used, by specific cleavage and ligation
CC of discontinuous nucleic acid, to generate new genes and gene products,
CC e.g. ribozymes (for use in gene therapy or as reagents in DNA
CC manipulation, e.g. replacements for restriction enzymes) or
CC immunologically active or signal-transducing proteins such as antibody
CC and enzyme libraries.
SQ Sequence 30 BP; 8 A; 8 C; 8 G; 6 T;

alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-3 x V66869
Align seg 1/1 to: V66869 from: 1 to: 30

seq_name: N_Geneseq_36:V66869
seq_documentation_block:
ID V66869 standard; cDNA; 34 BP.
AC V03294;
DT 22-JUN-1998 (first entry)
DE Isolate OC9a phosphatase (27A3A) gene 3' PCR primer.
KW Alkaline phosphatase; thermostable enzyme; thermophilic bacterium;
KW food; detergent; baking; PCR; primer; ss.
OS Synthetic.
OS Bacterial isolate OC9a-27A3A.
PN WO9748416-A1.
PD 24-DEC-1997.
PR 19-JUN-1996; US-033752.
PR 19-JUN-1996; US-033752.
PA (RECO-) RECOMBINANT BIOCATALYSIS INC.
PI Bylina E, Lee E, Mathur EJ;
DR WPI; 98-062851/06.
PT Thermostable phosphatase(s) - useful in pharmaceutical, food,
PT detergent, and baking industries
PS Example 1; Page 45; 128pp; English.
CC This 3' primer was used with a 5' primer (see V03293) to amplify
CC of bacterial isolate OC9a. The DNA was amplified from a pBluescript
CC vector, and was inserted into vector pGEM3. The invention relates
CC to claimed polynucleotides (see V03301-09 and V03310-20) coding for
CC claimed thermostable phosphatases (see W42380-95) that can be

```

CC utilised in the pharmaceutical, food, detergent, and baking
 CC industries.
 SQ Sequence 34 BP; 5 A; 7 C; 13 G; 9 T;

alignment_scores:
 Quality: 19.00 Length: 5
 Ratio: 3.800 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-08-653-294-3 x V03294/rev ..

Align seg 1/1 to reverse of: V03294 from: 1 to: 34

1 ArgGlu***LeuArg 5
 |||||

23 CGAGAGCCTTAAGG 9

seq_name: N_Geneseq_36:V82702

seq_documentation_block:
 ID V82702 standard; DNA: 34 BP.

AC V82702;
 DT 16-FEB-1999 (first entry)
 DE Antisense PCR primer C2y used amplify nucleic acid from FHBV variants.
 KW Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
 KW HBV-related disease; PCR primer; ss.

OS Synthetic.

OS Hepatitis B virus.

PN WO9845421-A2.

PD 15-OCT-1998.

PF 08-APR-1998; E02048.

PR 09-APR-1997; GB-007221.

PA (UNIU) UNIV GLASGOW.

PI Carman B;

DR WPI; 99-009329/01.

PT New hepatitis B virus nucleic acid with combination of specific

PT mutations - useful for, e.g. detection of binding interactions

PT between host or viral proteins and HBV nucleic

PS Disclosure; Page 52; 85pp; English.

CC PCR primers V82698-705 were used to amplify nucleic acid sequences
 CC from fulminant hepatitis B virus (FHBV) variants, in the course of
 CC the invention. The specification describes Hepatitis B virus (HBV)
 CC nucleic acid that has a mutation (i.e. alteration from the normal
 CC nucleotide in any of the genotypes A to F) in at least two of the
 CC enhancer I region, the negative regulatory element region, the enhancer
 CC II/ core upstream regulatory sequence/ basal core promoter region, or a
 CC mutation which leads to an X-peptide amino acid change to Cys or Met.

CC The HBV variants of the invention are used to detect binding interactions
 CC between host or viral proteins and HBV nucleic acid. Probes that
 CC hybridise to any of the specified mutated regions are used to detect

CC HBV-related disease, especially fulminant infection, but also severe
 CC chronic infection or serologically unusual forms of disease. Combinations

CC of the specified mutations are associated with fulminant infections,
 CC probably because they reduce the ability to bind inhibitory proteins in
 CC the host cell.

alignment_scores:
 Quality: 19.00 Length: 5
 Ratio: 3.800 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-08-653-294-3 x V82702/rev ..

Align seg 1/1 to reverse of: V82702 from: 1 to: 34

1 ArgGlu***LeuArg 5
 .|||||

24 CGGGAACCTTACGT 15

seq_name: N_Geneseq_36:T30827

seq_documentation_block:

ID T30827 standard; DNA; 41 BP.

AC T30827;

DT 12-SEP-1996 (first entry)

DE XhoI insertion primer XSVSD4647F.

KW Alphavirus; Sindbis virus; vector; gene therapy; vaccine;

KW primer; polymerase chain reaction; ELVIS; PCR; XhoI; SV40; ss.

OS Synthetic.

PN WO9617072-A2.

PD 06-JUN-1996.

PF 30-NOV-1995; UI5490.

PR 30-NOV-1994; US-348472.

PR 18-JAN-1995; US-376184.

PR 15-MAR-1995; US-405827.

PA (CHIR) CHIRON VIAGENE INC.

PI Belli BA, Chang SMW, Driver DA, Dubensky TW, Ibanez CE;

PI Jolly DJ, Polo JM;

DR WPI; 96-277785/28.

PT New recombinant alpha-virus vectors - used to develop prods and

PT methods for use in gene therapy and in the prodn. of vaccines

PS Example 3; Page 92; 256pp; English.

CC Primer XSVSD4647F (T30827) contains a buffer sequence allowing

CC enzyme digestion, a XhoI site, and nucleotides 4647-4675 of SV40.

CC It was used with reverse primer XSVSA562R (T30828) to insert an XhoI

CC site into the SV40 small t antigen intron sequence, using pBR322/SV40

CC (ATCC 45019) as template. This allowed the small t antigen intron to

CC be inserted into eukaryotic layered vector initiation systems (ELVIS)

CC (see also T30787) at a unique XhoI site immediately downstream of 5'

CC Sindbis sequences. Recombinant alphavirus expression vectors contg.

CC the SV40 promoter are useful in methods of gene therapy and for vaccine

CC prodn.

SQ Sequence 41 BP; 17 A; 6 C; 5 G; 13 T;

alignment_scores:

Quality: 19.00

Ratio: 3.800

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 5

Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 5

Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 5

Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 5

Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 5

Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 5

Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 5

Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 5

Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 5

Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 5

Gaps: 0

DR WPI; 98-446089/38.
PT DNA alpha-virus structural protein expression cassettes - for
PT producing recombinant alpha-virus particles
PS Example 3; Column 73; 140pp; English.
CC PCR primers V42367-420 and V42422-54 are used in the course of the
CC invention. The specification describes a DNA alphavirus structural
CC protein expression cassette which comprises an inducible promoter and
CC an alphavirus structural protein gene, where the promoter directs the
CC expression of the alphavirus structural protein gene upon induction of
CC the promoter within a cell, and where prior to induction within the
CC cell, the expression cassette does not express sufficient quantities of
CC structural proteins to be cytotoxic to a BHK cell containing the
CC expression cassette. The products may be used to inhibit pathogens and
CC stimulate an immune response.
SQ Sequence 41 BP; 17 A; 6 C; 5 G; 13 T;

alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-3 x V42404 ..
Align seg 1/1 to: V42404 from: 1 to: 41

1 ArgGlu***LeuArg 5
|||||
10 CGAGAAGCTCTAAGG 24

seq_name: N_Geneseq_36:V60164

seq_documentation_block:
ID V60164 standard; DNA; 41 BP.
AC V60164;
DT 04-DEC-1998 (first entry)
DE Forward primer XSVSD4647F used to amplify SV40 small t antigen introns.
KW Eukaryotic layered vector initiation system; stimulate; immune response;
KW Sindbis; PCR primer; ss.
OS Synthetic.
PN US5814482-A.
PD 29-SEP-1998.
PF 30-OCT-1996; 739158.
PR 15-SEP-1993; US-122791.
PR 18-FEB-1994; US-198450.
PR 14-SEP-1994; WO-010469.
PR 30-NOV-1994; US-348472.
PR 18-JAN-1995; US-376184.
PR 30-OCT-1996; US-739158.
PA (DRIV/) DRIVER D A.
PA (DUBE/) DUBENSKY T W.
PA (JOLLY/) JOLLY D J.
PA (POLO/) POLO J M.
PI Driver DA, Dubensky TW, Jolly DJ, Polo JM;
PT WPI; 98-541753/46.
DR Alphavirus layered vector initiation system - containing eukaryotic
PT promoter and heterologous antigen coding sequence, useful for
PT stimulating immune response
PS Example 3; Column 77; 144pp; English.
CC PCR primers V60164-65 are used to amplify the SV40 small t antigen
CC intron sequences. The product is used in the course of the invention.
CC The specification describes an eukaryotic layered vector
CC initiation system, based on Sindbis. The eukaryotic layered vector
CC which initiates, in a susceptible target cell, 5' to 3' synthesis of
CC RNA from the viral cDNA. The RNA comprises a vector construct which
CC autonomously amplifies in the cell and expresses a heterologous nucleic
CC acid sequence which encodes an antigen or modified form that stimulates
CC an immune response within an animal. The system is useful for stimulating
CC an immune response to an antigen by introducing the vector into target
CC cells, preferably by infection in vivo, especially where the immune
CC response is a cell mediated, HLA class I-restricted or an HLA class

CC II-restricted immune response.
SQ Sequence 41 BP; 17 A; 6 C; 5 G; 13 T;

alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-3 x V60164 ..
Align seg 1/1 to: V60164 from: 1 to: 41

1 ArgGlu***LeuArg 5
|||||
10 CGAGAAGCTCTAAGG 24

seq_name: N_Geneseq_36:V70724

seq_documentation_block:
ID V70724 standard; DNA; 41 BP.
AC V70724;
DT 02-FEB-1999 (first entry)
DE Forward primer XSVSD4647F used to amplify SV40 small t antigen introns.
KW Alphavirus vector construct; gene therapy; PCR primer; ss.
OS Synthetic.
PN US5843723-A.
PD 01-DEC-1998.
PF 30-OCT-1996; 739167.
PR 30-MAR-1995; US-404796.
PR 15-SEP-1993; US-122791.
PR 18-FEB-1994; US-198450.
PR 30-NOV-1994; US-348472.
PR 20-JAN-1995; US-376184.
PR 30-OCT-1996; US-739167.
PA (CHIR) CHIRON CORP.
PI Belli BA, Chang SMW, Driver DA, Dubensky TW, Ibanez CE,
PI Jolly DJ, Polo JM;
DR WPI; 99-044581/04.
PT Alphavirus vectors constructs containing a 5' promoter of viral cDNA
PT by in vitro transcription - used in gene therapy
PS Example 3; Column 73; 140pp; English.
CC PCR primers V70724-25 are used to amplify the SV40 small t antigen
CC intron sequences. The amplified product is used in the production
CC of the alphavirus vector constructs of the invention. These constructs
CC comprise a promoter 5' of viral cDNA which initiates the synthesis of
CC RNA from the viral cDNA by in vitro transcription, followed by a
CC 5' sequence which initiates transcription of alphavirus RNA.
CC followed by a nucleotide sequence encoding alphavirus nonstructural
CC proteins, a viral junction region which has been inactivated such
CC that viral transcription of a subgenomic fragment is prevented, an
CC internal ribosome entry site or a sequence which promotes ribosome
CC read through between adjacent reading frames, and an alphavirus
CC RNA polymerase recognition sequence. The recombinant alphavirus
CC vectors can be used for gene therapy.
SQ Sequence 41 BP; 17 A; 6 C; 5 G; 13 T;

alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-3 x V70724 ..
Align seg 1/1 to: V70724 from: 1 to: 41

1 ArgGlu***LeuArg 5
|||||
10 CGAGAAGCTCTAAGG 24

seq_name: N_Geneseq_36:Q48012

seq_documentation_block:

ID Q48012 standard; DNA; 45 BP.

AC Q48012;

DT 13-APR-1994 (first entry)

DE Listeria monocytogenes DNA detection probe 4.

KW Nucleic acid hybridisation assay; probe; food poisoning; detection;

KW 16S ribosomal RNA; ss.

OS Synthetic.

PN J05219997-A.

PD 31-AUG-1993.

PF 04-FEB-1992; 047959.

PR 04-FEB-1992; JP-047959.

PA (TOAG) TOA GOSEI CHEM IND LTD.

DR WPI; 93-308349/39.

PT Listeria monocytogenes detecting probe for testing foodstuffs or

PT patients - comprises DNA or RNA nucleic acid probe in which probe

PT is hybridised with nucleic acid of strain to be detected

PS Claim 1; Page 2; 7pp; Japanese.

CC This DNA sequence, its corresponding RNA version or complementary

CC sequence is claimed for detection of L.monocytogenes. The probes can

CC rapidly and sensitively detect the microorganism in food samples and

CC in patients.

SQ Sequence 45 BP; 8 A; 15 C; 10 G; 12 T;

alignment_scores:

Quality: 19.00

Ratio: 3.800

Percent Similarity: 100.000

Length: 5

Gaps: 0

Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x Q48012/rev ..

Align seg 1/1 to reverse of: Q48012 from: 1 to: 45

1 ArgGlu***LeuArg 5

|||||

26 CGTGAAGCACTAAGG 12

seq_name: N_Geneseq_36:Q58576

seq_documentation_block:

ID Q58576 standard; RNA; 45 BP.

AC Q58576;

DT 13-APR-1994 (first entry)

DE Listeria monocytogenes RNA detection probe 4.

KW Nucleic acid hybridisation assay; probe; food poisoning; detection;

KW 16S ribosomal RNA; ss.

OS Synthetic.

PN J05219997-A.

PD 31-AUG-1993.

PF 04-FEB-1992; 047959.

PR 04-FEB-1992; JP-047959.

PA (TOAG) TOA GOSEI CHEM IND LTD.

DR WPI; 93-308349/39.

PT Listeria monocytogenes detecting probe for testing foodstuffs or

PT patients - comprises DNA or RNA nucleic acid probe in which probe

PT is hybridised with nucleic acid of strain to be detected

PS Claim 1; Page 2; 7pp; Japanese.

CC This RNA sequence, its corresponding DNA version or complementary

CC sequence is claimed for detection of L.monocytogenes. The probes can

CC rapidly and sensitively detect the microorganism in food samples and

CC in patients.

SQ Sequence 45 BP; 8 A; 15 C; 10 G; 12 U;

alignment_scores:

Quality: 19.00

Ratio: 3.800

Percent Similarity: 100.000

Length: 5

Gaps: 0

Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x Q58576/rev ..

Align seg 1/1 to reverse of: Q58576 from: 1 to: 45

1 ArgGlu***LeuArg 5

|||||

26 CGTGAAGCACTAAGG 12

seq_name: N_Geneseq_36:Q63760

seq_documentation_block:

ID Q63760 standard; cDNA to mRNA; 47 BP.

AC Q63760;

DT 11-JAN-1995 (first entry)

DE Mink growth hormone/A. oryzae neutral protease pre-sequence oligomer.

KW Probe; isolation; mink growth hormone; MGH; rat growth hormone;

KW pre-sequence; neutral protease; Aspergillus oryzae; ss.

OS Synthetic.

PN CA2103623-A.

PD 27-FEB-1994.

PF 09-AUG-1993; 103623.

PR 26-AUG-1992; JP-227507.

PA (KIRK) KIRKMAN CORP.

PI Harada Y, Nakano E, Tatsumi H, Umezumi M;

DR WPI; 94-159680/20.

PT Recombinant prodn of mink growth hormone - using a synthetic DNA

PT fragment pre-sequence for efficient prodn. of the growth hormone

PT in Saccharomyces

PS Example; Page 24; 32pp; English.

CC The sequences given in Q63756-60 are synthetic oligomers which were

CC used in the construction of a DNA fragment encoding the mature mink

CC growth hormone (MGH) to which has been added the pre-sequence of

CC the neutral protease of Aspergillus oryzae and a hex 2 recognition

CC sequence. The mature mink growth hormone (MGH) coding sequence was

CC isolated using the probe given in Q63754. This probe was based on the

CC sequence of rat growth hormone and lead to the isolation of a 700 bp

CC fragment. Replacement of the mink growth hormone pre-sequence by the

CC pre-sequence derived from the neutral protease of Aspergillus oryzae

CC allows effective production of MGH by recombinant microorganisms.

SQ Sequence 47 BP; 16 A; 10 C; 16 G; 5 T;

alignment_scores:

Quality: 19.00

Ratio: 3.800

Percent Similarity: 100.000

Length: 5

Gaps: 0

alignment_block:

US-08-653-294-3 x Q63760 ..

Align seg 1/1 to: Q63760 from: 1 to: 47

1 ArgGlu***LeuArg 5

|||||

4 CGGGAATCTCTTAGA 18

OM of: US-08-653-294-3 to: EST:* out_format : pfs

Date: Feb 8, 2000 4:02 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xip
-Q/cgml_1/USPTO_spool/US08653294/runat_04022000_160700_15770/app_query.fasta.1
-DB=EST -QMT=fastap -SUFFIX=st -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000
-GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blossum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294
-NCPU=6 -ICPU=3 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-3

Query length: 10

Database: EST:

Database sequences: 4538634

Database length: 1887831982

Search time (sec): 8553.360000

score_list:

| Sequence | Strd | Orig | zscore | EScore | Len | Documentation |
|--------------------|------|-------|--------|---------|-----|-----------------------------------|
| gb_est19:AA811948 | + | 19.00 | 115.32 | 1.2e+03 | 43 | AA811948 ob74e04.s1 NCI_CGAP_GCB |
| gb_est16:AA564874 | + | 19.00 | 114.64 | 1.3e+03 | 40 | AA564874 n01b02.s1 NCI_CGAP_Pr2 |
| gb_est8:W89758 | + | 19.00 | 113.42 | 1.5e+03 | 49 | W89758 mf66c12.r1 Soares mouse e |
| gb_est21:AA948106 | + | 19.00 | 113.23 | 1.5e+03 | 50 | AA948106 on51a04.s1 NCI_CGAP_Cob |
| gb_est8:AA008561 | + | 19.00 | 111.37 | 1.9e+03 | 61 | AA008561 mg86d11.r1 Soares mouse |
| gb_est5:H63854 | + | 19.00 | 110.35 | 2.2e+03 | 68 | H63854 yu47c11.r1 Soares fetal 1 |
| gb_est20:AA872655 | + | 19.00 | 109.82 | 2.4e+03 | 72 | AA872655 o11e09.s1 NCI_CGAP_GC4 |
| gb_est21:AA995195 | + | 19.00 | 109.82 | 2.4e+03 | 72 | AA995195 ou90c11.s1 NCI_CGAP_Kid |
| gb_est17:AA621180 | + | 19.00 | 108.95 | 2.7e+03 | 79 | AA621180 af65h10.s1 Soares_NHMF |
| gb_est26:AA352574 | + | 19.00 | 108.83 | 2.7e+03 | 80 | AA352574 q19c07.r1 NCI_CGAP_GC4 |
| gb_est7:W96900 | + | 19.00 | 107.83 | 3.1e+03 | 89 | W96900 mf88a09.r1 Soares mouse e |
| gb_est23:AF027930 | + | 19.00 | 107.83 | 3.1e+03 | 89 | AF027930 AF027930 Mouse cDNA (S |
| gb_est18:AA701141 | + | 19.00 | 107.63 | 3.1e+03 | 91 | AA701141 zg56f04.s1 Soares pinea |
| gb_est17:AA611878 | + | 19.00 | 107.42 | 3.2e+03 | 93 | AA611878 vo90a04.s1 Bairstead mou |
| gb_est17:AA623328 | + | 19.00 | 107.42 | 3.2e+03 | 93 | AA623328 v119g10.r1 Bairstead mou |
| gb_est4:H52028 | + | 19.00 | 107.32 | 3.3e+03 | 94 | H52028 yo44e01.r1 Soares adult b |
| gb_est23:AA1088953 | + | 19.00 | 107.32 | 3.3e+03 | 94 | AA1088953 ou86b05.s1 Soares_NSF |
| gb_est28:AA1523638 | + | 19.00 | 107.32 | 3.3e+03 | 94 | AA1523638 tg95e01.s1 NCI_CGAP_CLI |
| gb_est7:W85542 | + | 19.00 | 107.03 | 3.4e+03 | 97 | W85542 mf54h10.r1 Soares mouse e |
| gb_est20:AA853669 | + | 19.00 | 107.03 | 3.4e+03 | 97 | AA853669 NHTSCae07e04f1 Normal H |
| gb_est23:AA151325 | + | 19.00 | 107.03 | 3.4e+03 | 97 | AA151325 qc74d02.r1 Soares place |
| gb_est3:BI3805 | + | 19.00 | 106.93 | 3.4e+03 | 98 | BI3805 NGE-SI/TGF3 40 Brugia mal |
| gb_est19:AA798893 | + | 19.00 | 106.84 | 3.5e+03 | 99 | AA798893 vv94e03.r1 Stratagene m |
| gb_est6:W59970 | + | 19.00 | 106.74 | 3.5e+03 | 100 | W59970 TgESTzy03d03.r1 TGRH Tac |
| gb_est9:AA095306 | + | 19.00 | 106.56 | 3.6e+03 | 102 | AA095306 12420.seq.F Human feta |
| gb_est11:AA235633 | + | 19.00 | 106.56 | 3.6e+03 | 102 | AA235633 12420.seq.F Human feta |
| gb_est12:AA319533 | + | 19.00 | 106.47 | 3.6e+03 | 103 | AA319533 ES21772 Adrenal gland |
| gb_est27:AA1432485 | + | 19.00 | 106.47 | 3.6e+03 | 103 | AA1432485 th35f04.s1 NCI_CGAP_Pa |
| gb_est37:AA1965303 | + | 19.00 | 106.47 | 3.6e+03 | 103 | AA1965303 fc89e05.s1 Zebrafish w |
| gb_est1:AA262426 | + | 19.00 | 106.37 | 3.7e+03 | 104 | AA262426 CITBI-E1-2502P13.TR CI |
| gb_est3:X78023 | + | 19.00 | 106.37 | 3.7e+03 | 104 | X78023 SCESTSHE6 Saccharomyces |
| gb_est9:AA091547 | + | 19.00 | 106.37 | 3.7e+03 | 104 | AA091547 m0068.seq.F Human feta |
| gb_gss11:W1846 | + | 19.00 | 106.20 | 3.8e+03 | 106 | W1846 zb59a10.r1 Soares_parrath |
| gb_gss11:AA264176 | + | 19.00 | 106.20 | 3.8e+03 | 106 | AA264176 CITBI-E1-2509A2.TF CIT |
| gb_gss13:AA485458 | + | 19.00 | 105.77 | 4.0e+03 | 111 | AA485458 RPCI-11-295K15.TV RPCI |
| gb_est25:AA1263925 | + | 19.00 | 105.68 | 4.0e+03 | 112 | AA1263925 qk02b09.s1 NCI_CGAP_Ki |
| gb_est36:AA1891243 | + | 19.00 | 105.60 | 4.1e+03 | 113 | AA1891243 614020G04.xl 614 - roo |
| gb_gss3:R07627 | + | 19.00 | 105.60 | 4.1e+03 | 114 | R07627 CDC64 Cri du chat, exon |
| gb_est12:AA292083 | + | 19.00 | 105.52 | 4.2e+03 | 116 | AA292083 t246e2.s1 Soares ovar |
| gb_est14:AA396862 | + | 19.00 | 105.35 | 4.2e+03 | 116 | AA396862 mr42d04.r1 Life Tech m |
| gb_est10:AA193629 | + | 19.00 | 105.27 | 4.3e+03 | 117 | AA193629 zr42c02.r1 Soares_NHMF |
| gb_gss12:AA20129 | + | 19.00 | 105.11 | 4.3e+03 | 119 | AA20129 RPCI-11-203L8.TV RPCI- |

gb_est5:D69141 - 19.00 105.04 4.4e+03 120 ! D69141 CELK06508F Yuji Kohar
gb_est18:AA680032 - 19.00 104.96 4.4e+03 121 ! AA680032 ag52e10.s1 Gessler
gb_est36:AV199574 - 19.00 104.96 4.4e+03 121 ! AV199574 AV199574 Yuji Kohar

seq_name: gb_est19:AA811948

seq_documentation_block:

LOCUS AA811948 40 bp mRNA EST 13-FEB-1998
DEFINITION ob74e04.s1 NCI_CGAP_GCB Homo sapiens cDNA clone IMAGE:1337118 3'
similar to SW:Z133_HUMAN P52736 ZINC FINGER PROTEIN 133. [1] ;
RNA sequence.

ACCESSION AA811948

VERSION AA811948.1 GI:2881559

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 40)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP).

Tumor Gene Index

Unpublished (1997)

On Sep 12, 1996 this sequence version replaced gi:1400907.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,

Ph.D., Gerald Marti, M.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 1.

Location/Qualifiers

source

1. 40

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1337118"

/clone_lib="NCI_CGAP_GCB1"

/tissue_type="germinal center B cell"

/lab_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was prepared from human tonsillar cells enriched for

germinal center B cells by flow sorting (CD20+, IgD-),

provided by Dr. Louis M. Staudt (NCI), Dr. David Allman

(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was

primed with a Not I - oligo(dT) primer

[5'-TGTTACCAATCTGAAGGGAGCGCGCTCATTTTTTTTTTTT-

3']. Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT7T3 vector. Library

went through one round of normalization, and was

constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 10 a 15 c 10 g 5 t

ORIGIN

alignment_scores:

Quality: 19.00 Length: 5

Ratio: 3.800 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x AA811948

Align seg 1/1 to: AA811948 from: 1 to: 40

1 ArgGlu***LeuArg 5
|||||
7 CGAGAGCCCTACGT 21

seq_name: gb_est16:AA564874

seq_documentation_block:
LOCUS AA564874 43 bp mRNA 04-SEP-1997
DEFINITION n01b02.s1 NCI-CCAP_Pr21 Homo sapiens cDNA clone IMAGE:985035 3' similar to SW:RS9_RAT P29314 40S RIBOSOMAL PROTEIN S9. ;, mRNA sequence.
ACCESSION AA564874
VERSION AA564874.1 GI:2336513
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 43)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1332888.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 772 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES

Source
1. 43
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:985035"
/clone_lib="NCI-CCAP_Pr21"
/sex="male"
/tissue_type="normal prostate"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared with normal prostate bulk tissue, and was then primed with a Not I - oligo(GT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is not normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 8 a 9 c 16 g 10 t
ORIGIN
alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-3 x AA564874

Align seg 1/1 to: AA564874 from: 1 to: 43

1 ArgGlu***LeuArg 5
|||||
6 CGGAGGCGCTTGAGG 20

seq_name: gb_est8:W89758

seq_documentation_block:
LOCUS W89758 49 bp mRNA 12-SEP-1996
DEFINITION m166c12.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone IMAGE:419254 5' similar to TR:G881954 G881954 RNPL. ;, mRNA sequence.
ACCESSION W89758
VERSION W89758.1 GI:1542432
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 49)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1405065.
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouses@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:253806

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 1.

FEATURES

Source
1. 49
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:419254"
/clone_lib="Soares mouse embryo NbME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(GT) primer [5', TGTTACCAATCTCAAGTGGGCGGCGGAAATTTTTTTTTTTTTTTT T 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 14 a 11 c 12 g 12 t
ORIGIN

alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x W89758

Align seg 1/1 to: W89758 from: 1 to: 49

1 ArgGlu***LeuArg 5

|||||

13 AGGGAACCTCTCGT 27

seq_name: gb_est21:AA948106

seq_documentation_block:

LOCUS AA948106

DEFINITION on51a04.s1 NCI-CGAP.Co8 Homo sapiens cDNA clone IMAGE:1560174 3'

similar to gb:M58297 ZINC FINGER PROTEIN 42 (HUMAN);, mRNA

sequence.

ACCESSION AA948106

VERSION AA948106.1 GI:3109359

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 50)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT On Jan 19, 1998 this sequence version replaced gi:2150604.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher Moskalkuk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Insert Length: 1844 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1..50

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1560174"

/clone_lib="NCI-CGAP.Co8"

/tissue_type="adenocarcinoma"

/lab_host="DH10B"

/note="Organ: Colon; Vector: p7T3b-Pac (Pharmacia) with a

modified polylinker; 1st strand cDNA was prepared from

colon adenocarcinoma, and was then primed with a Not I -

oligo(dT) primer. Double-stranded cDNA was ligated to Eco

RI adaptors (Pharmacia), digested with Not I and cloned

into the Not I and Eco RI sites of the modified p7T3

vector. Library is normalized. Library was constructed by

Bento Soares and M. Fatima Bonaldo."

Bento Soares and M. Fatima Bonaldo."

Bento Soares and M. Fatima Bonaldo."

Bento Soares and M. Fatima Bonaldo."

Bento Soares and M. Fatima Bonaldo."

Bento Soares and M. Fatima Bonaldo."

Bento Soares and M. Fatima Bonaldo."

Bento Soares and M. Fatima Bonaldo."

Bento Soares and M. Fatima Bonaldo."

Bento Soares and M. Fatima Bonaldo."

Bento Soares and M. Fatima Bonaldo."

Bento Soares and M. Fatima Bonaldo."

Bento Soares and M. Fatima Bonaldo."

Bento Soares and M. Fatima Bonaldo."

Bento Soares and M. Fatima Bonaldo."

Align seg 1/1 to reverse of: AA948106 from: 1 to: 50

1 ArgGlu***LeuArg 5

|||||

29 CGAGAGCCCTTCGC 15

seq_name: gb_est8:AA008561

seq_documentation_block:

LOCUS AA008561

DEFINITION mg86d11.r1 Soares mouse embryo NDME13.5 14.5 Mus musculus cDNA

clone IMAGE:439893 5' similar to PIR:S52089 S52089 hypothetical

protein - chinook salmon ;, mRNA sequence.

ACCESSION AA008561

VERSION AA008561.1 GI:1464528

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 61)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT On Jan 25, 1995 this sequence version replaced gi:637806.

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:265229

Trace considered overall poor quality

Possible reversed clone: similarity on wrong strand

Seq primer: ETPRimer

High quality sequence stop: 1.

Location/Qualifiers

1..61

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:439893"

/clone_lib="Soares mouse embryo NDME13.5 14.5"

/sex="unknown"

/tissue_type="embryo"

/dev_stage="13.5-14.5dpc total fetus"

/lab_host="DH10B"

/note="Vector: p7T3b-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5,

TGTACCAATCTGAAGTGGAGCGCCGCGGAATTTTITTTTTTTTTTTTTTTT

T 3'], on equal amounts of mRNA from 2 13.5dpc and 2

14.5dpc embryos (total RNA provided by Minoru KO, Wayne

State Univ., from 2); double-stranded cDNA was ligated to

Eco RI adaptors (Pharmacia), digested with Not I and

cloned into the Not I and Eco RI sites of the modified

p7T3 vector. Library went through one round of

normalization, and was constructed by Bento Soares and

M.Fatima Bonaldo."

M.Fatima Bonaldo."

M.Fatima Bonaldo."

M.Fatima Bonaldo."

M.Fatima Bonaldo."

M.Fatima Bonaldo."

M.Fatima Bonaldo."

M.Fatima Bonaldo."

M.Fatima Bonaldo."

BASE COUNT 13 a 12 c 16 g 20 t

ORIGIN

alignment_scores:

Quality: 19.00

Ratio: 3.800

Length: 5

Gaps: 0

alignment_scores:

Quality: 19.00

Ratio: 3.800

Length: 5

Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

Percent Identity: 100.000

Percent Identity: 100.000

Percent Identity: 100.000

Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x AA948106/rev

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x AA008561/rev ..

Align seg 1/1 to reverse of: AA008561 from: 1 to: 61

1 ArgGlu***LeuArg 5

57 CGAGAACCCCTCAGA 43

seq_name: gb_est5:H63854

seq_documentation_block:

LOCUS H63854 68 bp mRNA EST 11-OCT-1995
DEFINITION yu47c11.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:229268 5' similar to gb:L05093 60S RIBOSOMAL PROTEIN L18A
(HUMAN); mRNA sequence.

ACCESSION H63854

VERSION H63854.1 GI:1018655

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 68)

AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le.N.,
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R.
and Marra,M.

TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)

MEDLINE 97044478

COMMENT

On May 8, 1995 this sequence version replaced gi:800181.

Contact: Wilson RK
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

Insert Size: 699

High quality sequence starts: 1

High quality sequence stops: 1

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Trace considered overall poor quality

Insert length: 699 Std Error: 0.00

Seq primer: M13RP1

High quality sequence stop: 1.

Location/Qualifiers

FEATURES

source

FEATURES

source

ORIGIN

alignment_scores:

Quality: 19.00 Length: 5

Ratio: 3.800 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x H63854

Align seg 1/1 to: H63854 from: 1 to: 68

1 ArgGlu***LeuArg 5

24 CGCGAATCNTTGGC 38

seq_name: gb_est20:AA872655

seq_documentation_block:

LOCUS AA872655 72 bp mRNA EST 13-MAY-1998
DEFINITION o11ie09.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1476232 3'
similar to gb:M58297 ZINC FINGER PROTEIN 42 (HUMAN); mRNA
sequence.

ACCESSION AA872655

VERSION AA872655.1 GI:2968095

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 72)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT

On Sep 12, 1996 this sequence version replaced gi:1407100.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/dbrrp/image/image.html

Trace considered overall poor quality

Insert length: 817 Std Error: 0.00

Seq primer: -40m3 fwd. ET from Amersham

High quality sequence stop: 1.

Location/Qualifiers

FEATURES

source

1..72

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1476232"

/clone_lib="NCI_CGAP_GC4"

/tissue_type="pooled germ cell tumors"

/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; 1st strand cDNA was prepared from 3 pooled

germ cell tumors, and was then primed with a Not I -

oligo(dT) primer. Double-stranded cDNA was ligated to Eco

RI adaptors (Pharmacia), digested with Not I and cloned

into the Not I and Eco RI sites of the modified pT73

vector. Library is normalized. Library was constructed by

Bento Soares and M. Fatima Bonaldo."

8 a 19 c 28 g 17 t

BASE COUNT

ORIGIN

BASE COUNT

11 a 27 c 12 g 13 t 5 others

```

Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-08-653-294-3 x AA995195/rev

Align seg 1/1 to reverse of: AA872655 from: 1 to: 72

1 ArgGlu***LeuArg 5
|||||
36 AGAGAGCCCTACGC 22

seq_name: gb_est21:AA995195

seq_documentation_block:
LOCUS AA995195 72 bp mRNA EST 27-AUG-1998
DEFINITION
090c11.s1 NCI-CGAP_K1d3 Homo sapiens cDNA clone IMAGE:1635092 3'
similar to gb:M58297 ZINC FINGER PROTEIN 42 (HUMAN);, mRNA
sequence.
ACCESSION AA995195
VERSION AA995195.1 GI:3181684
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 72)
AUTHORS Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
COMMENT On May 5, 1995 this sequence version replaced gi:797755.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 809 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..72
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1635092"
/clone_lib="NCI-CGAP_K1d3"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. mRNA
source: 2 pooled kidneys. Library went through one round
of normalization. Library constructed by Bento Soares and
M. Fatima Bonaldo."
BASE COUNT 10 a 19 c 28 g 15 t
ORIGIN

alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-3 x AA872655/rev

Align seg 1/1 to reverse of: AA872655 from: 1 to: 72

1 ArgGlu***LeuArg 5
|||||
36 AGAGAGCCCTACGC 22

seq_name: gb_est21:AA995195

seq_documentation_block:
LOCUS AA995195 72 bp mRNA EST 27-AUG-1998
DEFINITION
090c11.s1 NCI-CGAP_K1d3 Homo sapiens cDNA clone IMAGE:1635092 3'
similar to gb:M58297 ZINC FINGER PROTEIN 42 (HUMAN);, mRNA
sequence.
ACCESSION AA995195
VERSION AA995195.1 GI:3181684
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 72)
AUTHORS Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
COMMENT On May 5, 1995 this sequence version replaced gi:797755.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 809 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..72
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1635092"
/clone_lib="NCI-CGAP_K1d3"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. mRNA
source: 2 pooled kidneys. Library went through one round
of normalization. Library constructed by Bento Soares and
M. Fatima Bonaldo."
BASE COUNT 10 a 19 c 28 g 15 t
ORIGIN

alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-3 x AA995195/rev

Align seg 1/1 to reverse of: AA995195 from: 1 to: 72

1 ArgGlu***LeuArg 5
|||||
36 CGAGAGCCCTACGC 22

seq_name: gb_est17:AA621180

seq_documentation_block:
LOCUS AA621180 79 bp mRNA EST 02-MAR-1998
DEFINITION
af65h10.s1 Soares_NbHMPu_s1 Homo sapiens cDNA clone IMAGE:1046947
3' similar to gb:X07290_cds1 ZINC FINGER PROTEIN HF.12
(HUMAN); contains Alu repetitive element; mRNA sequence.
ACCESSION AA621180
VERSION AA621180.1 GI:2525119
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 79)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1407267.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 188 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..79
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1046947"
/clone_lib="Soares_NbHMPu_s1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NbHM, pregnant uterus
NbHPU, and fetal heart NbH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
BASE COUNT 16 a 20 c 18 g 25 t
ORIGIN

alignment_scores:
Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment_block:

US-08-653-294-3 x AA621180/rev

Align seg 1/1 to reverse of: AA621180 from: 1 to: 79

1 ArgGlu***LeuArg 5

|||||

40 AGAGAAACCTTCGA 26

seq_name: gb_est36:AI352574

seq_documentation_block:

LOCUS AI352574 80 bp mRNA EST 30-DEC-1998
 DEFINITION q19c07.x1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1948044 3',
 mRNA sequence.

ACCESSION AI352574

VERSION AI352574.1 GI:4089780

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 80)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP).

Tumor Gene Index

Unpublished (1997)

On Jan 17, 1998 this sequence version replaced gi:1900526.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Cloned through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40UP from Gibco

High quality sequence stop: 56.

Location/Qualifiers

FEATURES

source

1..80
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:1948044"
 /clone_lib="NCI_CGAP_GC4"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"

/note="vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from 3 pooled
 germ cell tumors, and was then primed with a Not I -
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pT7T3
 vector. Library is normalized. Library was constructed by
 Bento Soares and M. Fatima Bonaldo."

BASE COUNT

ORIGIN

26 a 10 c 15 g 29 t

alignment_scores:

Quality: 19.00 Length: 5
 Ratio: 3.800 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x AI352574

Align seg 1/1 to: AI352574 from: 1 to: 80

1 ArgGlu***LeuArg 5

..

|||||

39 AGAGAAGCCCTTAGA 53

seq_name: gb_est7:W96900

seq_documentation_block:

LOCUS W96900 89 bp mRNA EST 16-JUL-1996
 DEFINITION mf88a09.r1 Soares mouse embryo NbMEL3.5 14.5 Mus musculus CDNA
 clone IMAGE:421336 5' similar to gb:L35249 VACUOLAR ATP SYNTHASE
 SUBUNIT B, BRAIN ISOFORM (HUMAN);, mRNA sequence.

ACCESSION W96900

VERSION W96900.1 GI:1426887

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 89)

AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL

COMMENT

On May 18, 1995 this sequence version replaced gi:811252.

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:255888

Trace considered overall poor quality

Seq primer: EMP1mer

High quality sequence stop: 1.

FEATURES

source

1..89
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_image="IMAGE:421336"
 /clone_lib="Soares mouse embryo NbMEL3.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"

/note="vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I, Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TTTTACCATCTGAAGTGGAGCGCGGAAATTTTTTTTTTTTTTTTTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos (total RNA provided by Minoru Ko, Wayne
 State Univ., from 2 l); double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT7T3 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M.Fatima Bonaldo."

BASE COUNT

ORIGIN

27 a 24 c 19 g 19 t

alignment_scores:

Quality: 19.00 Length: 5
 Ratio: 3.800 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x W96900

JOURNAL
COMMENT

Unpublished (1996)
On Sep 12, 1996 this sequence version replaced gi:1402225.
Contact: Marra M/Mouse EST Project
WASHU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:588708

Trace considered overall poor quality
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers

FEATURES
source

1. .93
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:1066348"
/clone_lib="Barstead mouse irradiated colon MPLRB7"
/dev_stage="8 weeks"
/lab_host="DH10B"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
from 8 week old mouse. Colon was harvested 72 hours after
irradiation with 1400 Gys. 1st strand cDNA was primed
with a Not I - oligo (dT) primer
[5'TGTGAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors [AATTCGATCCTTG], digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library constructed by Bob Barstead."

BASE COUNT
ORIGIN

24 a 20 c 24 g 25 t

alignment_scores:

Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x AA611878/rev ..
Align seg 1/1 to reverse of: AA611878 from: 1 to: 93

1 ArgGlu***LeuArg 5

|||||
76 AGAGAGGCGCTTACGG 62

seq_name: gb_est17:AA623328

seq_documentation_block:

LOCUS AA623328 93 bp mRNA EST 14-OCT-1997
DEFINITION v19g10.r1 Barstead mouse proximal colon MPLRB6 Mus musculus cDNA
clone IMAGE:904290 5' similar to gb:X79535 TUBULIN BETA-2 CHAIN
(HUMAN); gb:N28732 Mouse beta-tubulin gene M-beta-5, 3' end
(MOUSE); mRNA sequence.

ACCESSION AA623328

VERSION AA623328.1 GI:2527204

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 93)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

TITLE
JOURNAL
COMMENT

The WashU-HHMI Mouse EST Project
Unpublished (1996)
On Nov 29, 1993 this sequence version replaced gi:636298.
Contact: Marra M/Mouse EST Project
WASHU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:524954

Trace considered overall poor quality
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers

FEATURES
source

1. .93
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:904290"
/clone_lib="Barstead mouse proximal colon MPLRB6"
/dev_stage="7 day juvenile"
/lab_host="DH10B"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
was primed with a Not I - oligo (dT) primer [5'
TGTACGAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[AATTCGATCCTTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT7T3 vector.
Library constructed by Bob Barstead."

BASE COUNT 27 a 21 c 34 g 11 t

alignment_scores:

Quality: 19.00 Length: 5
Ratio: 3.800 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-3 x AA623328 ..
Align seg 1/1 to: AA623328 from: 1 to: 93

1 ArgGlu***LeuArg 5

|||||
60 AGGGAGAGTTTGAGG 74

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: February 8, 2000, 01:29:35 ; Search time 122.56 Seconds
(without alignments)
1.160 Million cell updates/sec

Title: US-08-653-294-4

Perfect score: 29

Sequence: 1 RIALRY 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|------------------------|
| 1 | 29 | 100.0 | 6 | 1 W47261 | Immunomodulatory p |
| 2 | 29 | 100.0 | 6 | 1 W33780 | Peptide #1 used in |
| 3 | 29 | 100.0 | 10 | 1 R41208 | Peptide fragment o |
| 4 | 29 | 100.0 | 10 | 1 R33062 | HLA-B*2702 CTL modu |
| 5 | 29 | 100.0 | 10 | 1 R33094 | HLA-B*2702 CTL modu |
| 6 | 29 | 100.0 | 10 | 1 R35413 | Alpha1-helix of HL |
| 7 | 29 | 100.0 | 10 | 1 R35425 | HLA-B*2702.75-84(D) |
| 8 | 29 | 100.0 | 10 | 1 W07512 | T-cell modulating |
| 9 | 29 | 100.0 | 10 | 1 W07513 | T-cell modulating |
| 10 | 29 | 100.0 | 10 | 1 W47265 | Immunomodulatory p |
| 11 | 29 | 100.0 | 10 | 1 W47269 | Immunomodulatory p |
| 12 | 29 | 100.0 | 10 | 1 W33784 | Peptide B2702.75-8 |
| 13 | 29 | 100.0 | 10 | 1 W33787 | Peptide B2702.75-8 |
| 14 | 29 | 100.0 | 12 | 1 R35429 | HLA-B*2702.84-79-84 |
| 15 | 29 | 100.0 | 12 | 1 W33798 | Peptide B2702.84-7 |
| 16 | 29 | 100.0 | 12 | 1 W33799 | Immunomodulating d |
| 17 | 29 | 100.0 | 15 | 1 R32912 | HLA-B*2702 CTL modu |
| 18 | 29 | 100.0 | 15 | 1 W33795 | Peptide B2702.70-8 |
| 19 | 29 | 100.0 | 20 | 1 R32907 | HLA-B*2702 CTL modu |
| 20 | 29 | 100.0 | 20 | 1 R32908 | HLA-B*2702 CTL modu |
| 21 | 29 | 100.0 | 20 | 1 R35428 | HLA-B*2702.84-75-84 |
| 22 | 29 | 100.0 | 20 | 1 W33778 | Immunomodulating d |
| 23 | 29 | 100.0 | 20 | 1 W33791 | Peptide B2702.84-7 |
| 24 | 29 | 100.0 | 25 | 1 R41205 | Peptide fragment o |
| 25 | 29 | 100.0 | 25 | 1 R48286 | Peptide fragment o |
| 26 | 29 | 100.0 | 25 | 1 R33090 | HLA-B*2702 CTL modu |
| 27 | 29 | 100.0 | 25 | 1 R33093 | HLA-B*2702 CTL modulat |
| 28 | 29 | 100.0 | 25 | 1 R35416 | HLA-B*2702.60-84. C |
| 29 | 29 | 100.0 | 25 | 1 R35422 | HLA-B*38.6084. Comps |
| 30 | 29 | 100.0 | 25 | 1 R33794 | Peptide B2702.60-8 |
| 31 | 29 | 100.0 | 184 | 1 Y06801 | Peptide Seq ID No: |
| 32 | 29 | 100.0 | 362 | 1 R03142 | Sequence of HLA-B*5 |
| 33 | 29 | 100.0 | 362 | 1 R03144 | Sequence of HLA-B*5 |
| 34 | 29 | 100.0 | 362 | 1 R12463 | HLA-B*53 exon. HLA |

ALIGNMENTS

RESULT 1

W47261
ID W47261 standard; peptide; 6 AA.
AC W47261;
DT 22-MAY-1998 (first entry)
DE Immunomodulatory peptide.
KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
KW transplant rejection; treatment; autoimmune disease.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 1..6 /note= "at least one of the amino acids is the D-isomer"
FT PN W09744052-A1.
PD 27-NOV-1997.
PF 23-APR-1997; U06705.
PR 22-MAY-1996; US-651650.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM;
DR WPI; 98-018220/02.
PT Novel immunomodulatory peptide-type compound - useful for inhibiting transplant rejection
PS Claim 10; Page 36; 41pp; English.
CC The present sequence is an immunomodulatory peptide, which comprises a Class I HLA-B alpha-1 domain sequence. It can be used in a pharmaceutical composition together with a subtherapeutic dose of an immunosuppressant, to extend the period of acceptance of a transplant from a major histocompatibility complex (MHC) unmatched donor, i.e. to inhibit transplant rejection. It can also be used in the treatment of autoimmune diseases.
CC Peptides using the D-form amino acids are more effective
CC immunomodulators than their diastereomers or enantiomers.
SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RIALRY 6
| | | | |

Db 1 RIALRY 6

RESULT 2

W33780
ID W33780 standard; peptide; 6 AA.
AC W33780;
DT 19-JUN-1998 (first entry)
DE Peptide #1 used in immunomodulating dimer peptide.
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
KW transplacation; autoimmune disease; Class I HLA-B alpha-1 domain; rejection.
OS Synthetic.
OS Homo sapiens.
PN W09744351-A1.

35 28 96.6 337 1 W09110 Human amine recept
36 27 93.1 386 1 W99399 S. antibiotic ole
37 27 93.1 547 1 W46202 Rice RSW1-like cel
38 27 93.1 1065 1 W33819 Arabidopsis cellul
39 26 89.7 141 1 W38584 S. pneumoniae pfs
40 26 89.7 3567 1 R44431 eryA region polype
41 25 86.2 939 1 R04104 pJH1 gene product
42 25 86.2 959 1 R76066 Yeast MSH1 protein
43 25 86.2 971 1 R76070 Yeast MSH1 protein
44 25 86.2 1081 1 W33817 Arabidopsis cellul
45 25 86.2 1081 1 W33820 Arabidopsis cellul

PD 27-NOV-1997. U08689.
 PF 22-MAY-1997; U08689.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI; 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Claim 15; Page 35; 41pp; English.
 CC This sequence represents a specifically claimed peptide which forms part
 of the immunomodulating dimer peptides of the invention. A peptide-type
 compound or variant is claimed which has immunomodulating activity,
 CC including the N-terminal acylated and/or C-terminal amidated or
 CC esterified forms of up to 60 amino acids, where the peptide-type compound
 CC comprises the formula: A-B, where A, B = (R aa76-77L) (aa79-84) or
 CC (aa80-79) (Laa77-76R); aa76 = E or V; aa77 = D, S or N; aa79 = R or G;
 CC aa80 = I or N; aa81, aa84 = a hydrophobic or small amino acid; aa82 = R
 CC or L; aa83 = G or R; and aa represents amino acid. The sequence in the
 CC brackets may optionally be absent or truncated at any peptide type bond
 CC within the brackets. The compounds comprise amino acid sequences related
 CC to a Class I HLA-B alpha domain (positions 79-84). They can be used to
 CC inhibit cytotoxic T-lymphocytes (CTL) from undesirably attacking cells in
 CC a host or in vitro. They can also be used in combination with antigenic
 CC peptides or proteins of interest to activate CTLs. They can also inhibit
 CC the proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.
 SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0;

QY 1 RIALRY 6
 |||||
 DB 1 RIALRY 6

RESULT 3

R41208
 ID R41208 standard; peptide; 10 AA.
 AC R41208;
 DT 15-MAR-1994 (first entry)
 DE Peptide fragment of Class I HLA peptide.
 KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;
 KW parasitic disease; cytotoxic T lymphocyte; modulation.
 OS Synthetic.
 PN W09317695-A.
 PD 16-SEP-1993.
 PF 23-FEB-1993; U01758.
 PR 02-MAR-1992; US-844716.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger CA, Krensky AM;
 DR WPI; 93-303134/38.
 PT New peptide(s) based on Class I HLA antigen domains - used for
 PT modulating cytotoxic T-lymphocyte activity towards targets
 PS Claim 11; Page 54; 61pp; English.
 CC The peptide is used to modulate cytotoxic T-lymphocyte (CTL)
 CC activity, either by inhibition or stimulation. It can be used
 CC for inhibiting CTL toxicity in transplantations, for inducing CTL
 CC activity in parasitic diseases and neoplasia and in studies on viral
 CC infection. The peptide can also be used for identifying CTLs which
 CC bind to it and removing subsets of CTLs from a T-cell composition.
 CC This peptide sequence is more commonly found within larger peptide
 CC compounds of not more than 30 amino acids in length.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.23;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RIALRY 6
 |||||
 DB 5 RIALRY 10

RESULT 4

R83062
 ID R83062 standard; peptide; 10 AA.
 AC R83062;
 DT 16-MAY-1996 (first entry)
 DE HLA-B2702 CTL modulating peptide (B2702.75-84).
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW Class I MHC; HLA-B2702.
 OS Synthetic.
 PN W09526979-A1.
 PD 12-OCT-1995.
 PF 05-APR-1995; U04349.
 PR 05-APR-1994; US-222851.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM, Parham P;
 DR WPI; 95-358582/46.
 PT Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 PT host
 PS Claim 15; Page 9; 80pp; English.
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of
 CC class I major histocompatibility complex (MHC) antigens. This sequence
 CC corresponds to residues 75-84 of the alpha-1 domain of the class I MHC
 CC HLA-B2702. These sequences can be used to extend the period of
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.
 CC The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.23;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 |||||
 DB 5 RIALRY 10

RESULT 5

R83094
 ID R83094 standard; peptide; 10 AA.
 AC R83094;
 DT 16-MAY-1996 (first entry)
 DE HLA-B2702 CTL modulating peptide (B2702.75-84(D)).
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW Class I MHC; HLA-B2702.
 OS Synthetic.
 PN W09526979-A1.
 PD 12-OCT-1995.
 PF 05-APR-1995; U04349.
 PR 05-APR-1994; US-222851.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM, Parham P;
 DR WPI; 95-358582/46.
 PT Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 PT host
 PS Example 14; Page 34; 80pp; English.
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of

CC class I major histocompatibility complex (MHC) antigens. This sequence
 CC corresponds to residues 75-84 of the alpha-1 domain of the class I MHC
 CC HLA-B*2702. These sequences can be used to extend the period of
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.
 CC The peptides are administered to a patient in conjunction with
 CC a subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 CC Sequence 10 AA;
 SQ

Query Match 100.0%; Score 29; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 |||||
 DB 5 RIALRY 10

RESULT 6

ID R95413 standard; peptide; 10 AA.
 AC R95413;
 DT 12-NOV-1996 (first entry)
 DE Alpha-helix of HLA-B*2702.
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;
 CC T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 CC cytolysis; antigen presenting cell.
 OS Synthetic.
 PN W09513288-Al.
 PD 18-MAY-1995.
 PF 10-NOV-1994; U12985.
 PR 10-NOV-1993; US-150493.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI: 95-194027/25.
 PT Compsns. comprising lymphoid surface membrane proteins - which may
 PT inhibit cytolytic activity and differentiation of CTLs.
 PS Example; Page 11; 29pp; English.

CC This sequence represents the alpha-helix of the
 CC human-leucocyte-associated antigen B2702 (HLA-B*2702). This sequence,
 CC epitopes, and palindromes of it (such as R95428) can be used to isolate
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
 CC protein associated with T-cell activation in mammalian T-cells, and is
 CC also immunologically cross reactive with the heat shock protein Hsc70.
 CC p74 is found in a limited number of cell types, but is particularly
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable
 CC cell with an amphoteric detergent, and then passed through an affinity
 CC column containing a covalently bound HLA-B*2702 palindromic peptide.
 CC Compositions comprising the extracellular fragment of p74 combined with
 CC HLA-B*2702.60-84 (see R95416), induces calcium influx, and inhibits
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate
 CC compounds can be screened for their effect on the cytolytic activity of
 CC T-cells, by combining them with the extracellular portion of p74 and
 CC determining the amount of binding between the candidate compound and p74.
 CC Modulation of CTL activity can be inhibited in a cellular composition
 CC containing T-cells and antigen presenting cells (APCs), by adding to the
 CC mix the extracellular portion of p74, in an amount sufficient to compete
 CC with p74 for the binding of the p74 ligand.
 CC Sequence 10 AA;
 SQ

Query Match 100.0%; Score 29; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 |||||
 DB 5 RIALRY 10

RESULT 7

ID R95425 standard; peptide; 10 AA.
 AC R95425;
 DT 12-NOV-1996 (first entry)
 DE HLA-B*2702.75-84(D).
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;
 CC T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 CC cytolysis; antigen presenting cell.
 OS Synthetic.
 FH Key
 FT misc_difference 3 Location/Qualifiers
 FT /note= "N3D mutation"

FN W09513288-Al.
 PD 18-MAY-1995.
 PF 10-NOV-1994; U12985.
 PR 10-NOV-1993; US-150493.

PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI: 95-194027/25.

PT Compsns. comprising lymphoid surface membrane proteins - which may
 PT inhibit cytolytic activity and differentiation of CTLs.
 PS Example; Page 11; 29pp; English.
 CC R95413, and R95415-R95431 represent palindromes and fragments of
 CC human-leucocyte-associated antigens. This sequence represents the
 CC HLA-B*2702.75-84(D). These sequences can be used to isolate the protein
 CC p74 from a T-cell lysate. p74 is a T-cell surface membrane protein
 CC associated with T-cell activation in mammalian T-cells, and is also
 CC immunologically cross reactive with the heat shock protein Hsc70. p74 is
 CC found in a limited number of cell types, but is particularly expressed on
 CC B and T cells. p74 can be isolated by lysis of a suitable cell with an
 CC amphoteric detergent, and then passed through an affinity column
 CC containing a covalently bound HLA-B*2702 palindromic peptide.
 CC Compositions comprising the extracellular fragment of p74 combined with
 CC HLA-B*2702.60-84 (see R95416), induces calcium influx, and inhibits
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate
 CC compounds can be screened for their effect on the cytolytic activity of
 CC T-cells, by combining them with the extracellular portion of p74 and
 CC determining the amount of binding between the candidate compound and p74.
 CC Modulation of CTL activity can be inhibited in a cellular composition
 CC containing T-cells and antigen presenting cells (APCs), by adding to the
 CC mix the extracellular portion of p74, in an amount sufficient to compete
 CC with p74 for the binding of the p74 ligand.
 CC Sequence 10 AA;
 SQ

Query Match 100.0%; Score 29; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 |||||
 DB 5 RIALRY 10

RESULT 8

ID W07512 standard; peptide; 10 AA.

AC W07512;

DT 04-AUG-1997 (first entry)

DE T-cell modulating peptide #1.

KW mammal; major histocompatibility complex; MHC class I; antigen; perforin;
 KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;
 KW rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;
 KW thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;
 KW autologous target cell; cytokine release; T cell activation; therapy.
 OS Synthetic.
 PN W09635443-Al.
 PD 14-NOV-1996.

PF 05-APR-1996; U04710.
 PR 12-MAY-1995; US-440504.
 PA (SANG-) SANGSTAT MEDICAL CORP.
 PI Buelow R;
 DR WPI: 96-518410/51.
 PT Treatment of auto-immune disease by admin. of peptide(s) corresp. to
 PT major histocompatibility complex antigens - esp. for delaying onset
 PT of clinical symptoms of insulin dependent diabetes by modulating T
 PT cell mediated attack on target cells
 PS Claim 7; Page 20; 24pp; English.
 CC W07512-W07518 represent T-cell modulating peptides that can be used in
 CC the method of the invention. These sequences are based on a portion of
 CC the generic peptide corresponding to residues 70-91 of the alpha-1 domain
 CC of the major histocompatibility complex (MHC) class I antigen (see
 CC W07510). The method is for affecting the course of an autoimmune disease
 CC involving T-cell mediated destruction of tissue in mammals. These
 CC peptides are used especially to treat insulin-dependent diabetes
 CC mellitus, preferably being administered during the pre-clinical stage to
 CC delay onset of the disease. Other diseases that can be treated are
 CC multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus vulgaris,
 CC Sjogren's disease, thyroid disease, Hashimoto's thyroiditis, myasthenia
 CC gravis, etc. The peptides modulate T-cell mediated attack on autologous
 CC target cells, and may also reduce inflammation, swelling, and release of
 CC cytokines, perforins, granzymes etc. associated with T cell activation.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 Db 5 RIALRY 10

RESULT 9
 W07513
 ID W07513 standard; peptide; 10 AA.
 AC W07513; 1997 (first entry)
 DE T-cell modulating peptide #2.
 KW T-cell modulator; autoimmune disease; tissue destruction; alpha-1 domain;
 KW mammal; major histocompatibility complex; MHC class I; antigen; perforin;
 KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;
 KW rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;
 KW thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;
 KW autologous target cell; cytokine release; T cell activation; therapy.
 OS Synthetic.
 PN W0963543-A1.
 PD 14-NOV-1996.
 PF 05-APR-1996; U04710.
 PR 12-MAY-1995; US-440504.
 PA (SANG-) SANGSTAT MEDICAL CORP.
 PI Buelow R;
 DR WPI: 96-518410/51.
 PT Treatment of auto-immune disease by admin. of peptide(s) corresp. to
 PT major histocompatibility complex antigens - esp. for delaying onset
 PT of clinical symptoms of insulin dependent diabetes by modulating T
 PT cell mediated attack on target cells
 PS Claim 7; Page 20; 24pp; English.
 CC W07512-W07518 represent T-cell modulating peptides that can be used in
 CC the method of the invention. These sequences are based on a portion of
 CC the generic peptide corresponding to residues 70-91 of the alpha-1 domain
 CC of the major histocompatibility complex (MHC) class I antigen (see
 CC W07510). The method is for affecting the course of an autoimmune disease
 CC involving T-cell mediated destruction of tissue in mammals. These
 CC peptides are used especially to treat insulin-dependent diabetes
 CC mellitus, preferably being administered during the pre-clinical stage to
 CC delay onset of the disease. Other diseases that can be treated are
 CC multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus vulgaris,
 CC Sjogren's disease, thyroid disease, Hashimoto's thyroiditis, myasthenia
 CC gravis, etc. The peptides modulate T-cell mediated attack on autologous

CC target cells, and may also reduce inflammation, swelling, and release of
 CC cytokines, perforins, granzymes etc. associated with T cell activation.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 Db 5 RIALRY 10

RESULT 11
 W47269
 ID W47269 standard; peptide; 10 AA.
 AC W47269;
 DT 22-MAY-1998 (first entry)
 DE Immunomodulatory peptide.
 KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
 KW transplant rejection; treatment; autoimmune disease.
 OS Homo sapiens.
 FH Key
 FT Misc_difference 1.10
 FT /note- "at least one of the amino acids is the D-isomer

CC target cells, and may also reduce inflammation, swelling, and release of
 CC cytokines, perforins, granzymes etc. associated with T cell activation.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 Db 5 RIALRY 10

RESULT 10
 W47265
 ID W47265 standard; peptide; 10 AA.
 AC W47265;
 DT 22-MAY-1998 (first entry)
 DE Immunomodulatory peptide.
 KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
 KW transplant rejection; treatment; autoimmune disease.
 OS Homo sapiens.
 FH Key
 FT Misc_difference 1.10
 FT /note- "at least one of the amino acids is the D-isomer

CC target cells, and may also reduce inflammation, swelling, and release of
 CC cytokines, perforins, granzymes etc. associated with T cell activation.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 Db 5 RIALRY 10

RESULT 11
 W47269
 ID W47269 standard; peptide; 10 AA.
 AC W47269;
 DT 22-MAY-1998 (first entry)
 DE Immunomodulatory peptide.
 KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
 KW transplant rejection; treatment; autoimmune disease.
 OS Homo sapiens.
 FH Key
 FT Misc_difference 1.10
 FT /note- "at least one of the amino acids is the D-isomer

PN WO9744052-A1.
PD 27-NOV-1997.
PF 23-APR-1997; U06705.
PR 22-MAY-1996; US-651650.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM;
DR WPI: 98-018220/02.
PT Novel immunomodulatory peptide-type compound - useful for inhibiting
transplant rejection
PS Claim 10; Page 36; 41pp; English.
CC The present sequence is an immunomodulatory peptide, which
comprises a Class I HLA-B alpha-1 domain sequence. It can be used
in a pharmaceutical composition together with a subtherapeutic dose
of an immunosuppressant, to extend the period of acceptance of a
transplant from a major histocompatibility complex (MHC) unmatched
donor, i.e. to inhibit transplant rejection. It can also be used in
the treatment of autoimmune diseases.
CC Peptides using the D-form amino acids are more effective
immunomodulators than their diastereomers or enantiomers.
SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RIALRY 6
DB 5 RIALRY 10
RESULT 12
W33784
ID W33784 standard; peptide; 10 AA.
AC W33784;
DT 19-JUN-1998 (first entry)
DE Peptide B2702.75-84 tested for immunomodulating activity.
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
KW rejection.
OS Synthetic.
OS Homo sapiens.
PN WO9744351-A1.
PD 27-NOV-1997.
PF 22-MAY-1997; U06689.
PR 24-MAY-1996; US-653294.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Beulow R, Clayberger C, Krensky AM;
DR WPI: 98-086530/08.
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
alpha-1 domain, used for preventing rejection of transplants or
treating autoimmune diseases
PS Example 1; Page 19; 41pp; English.
CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
activity. A peptide-type compound or variant is claimed which has
immunomodulating activity, including the N-terminal acylated and/or
C-terminal amidated or esterified forms of up to 60 amino acids, where
the peptide-type compound comprises the formula: A-B, where A, B =
(R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
acid. The sequence in the brackets may optionally be absent or truncated
at any peptide type bond within the brackets. The compounds comprise
amino acid sequences related to a Class I HLA-B alpha domain (positions
79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
undesirably attacking cells in a host or in vitro. They can also be
used in combination with antigenic peptides or proteins of interest to
activate CTLs. They can also inhibit the proliferation of T cells in
transplants or for treating autoimmune diseases, e.g. diabetes,
rheumatoid arthritis and lupus erythematosus. The products can also be
used for detection and diagnosis.
SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RIALRY 6
DB 5 RIALRY 10
RESULT 14
R95429
ID R95429 standard; peptide; 12 AA.
AC R95429;
DT 12-NOV-1996 (first entry)
DE HLA-B*2702.84-79-84 palindromic.
KW HLA; p4; alpha1-helix; human-leucocyte-associated antigen; inhibitor;
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RIALRY 6
DB 5 RIALRY 10
RESULT 13
W33787
ID W33787 standard; peptide; 10 AA.
AC W33787;
DT 19-JUN-1998 (first entry)
DE Peptide B2702.75-84p77 tested for immunomodulating activity.
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
KW rejection.
OS Synthetic.
OS Homo sapiens.
PN WO9744351-A1.
PD 27-NOV-1997.
PF 22-MAY-1997; U08689.
PR 24-MAY-1996; US-653294.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Beulow R, Clayberger C, Krensky AM;
DR WPI: 98-086530/08.
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
alpha-1 domain, used for preventing rejection of transplants or
treating autoimmune diseases
PS Example 1; Page 19; 41pp; English.
CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
activity. A peptide-type compound or variant is claimed which has
immunomodulating activity, including the N-terminal acylated and/or
C-terminal amidated or esterified forms of up to 60 amino acids, where
the peptide-type compound comprises the formula: A-B, where A, B =
(R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
acid. The sequence in the brackets may optionally be absent or truncated
at any peptide type bond within the brackets. The compounds comprise
amino acid sequences related to a Class I HLA-B alpha domain (positions
79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
undesirably attacking cells in a host or in vitro. They can also be
used in combination with antigenic peptides or proteins of interest to
activate CTLs. They can also inhibit the proliferation of T cells in
transplants or for treating autoimmune diseases, e.g. diabetes,
rheumatoid arthritis and lupus erythematosus. The products can also be
used for detection and diagnosis.
SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RIALRY 6
DB 5 RIALRY 10
RESULT 14
R95429
ID R95429 standard; peptide; 12 AA.
AC R95429;
DT 12-NOV-1996 (first entry)
DE HLA-B*2702.84-79-84 palindromic.
KW HLA; p4; alpha1-helix; human-leucocyte-associated antigen; inhibitor;
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
SQ Sequence 10 AA;

KW cytotoxic; antigen presenting cell.
OS Synthetic.
PN WO9513288-A1.
PD 18-MAY-1995.
PF 10-NOV-1994; U12985.
PR 10-NOV-1993; US-150493.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM;
DR WPI: 95-194027/25.
PT Compens. comprising lymphoid surface membrane proteins - which may
inhibit cytolytic activity and differentiation of CTLs.
PS Example; Page 12: 29pp; English.
CC R95413, and R95415-R95431 represent palindromes and fragments of
human-leucocyte-associated antigens. This sequence represents the
HLA-B2702 84-79-84 palindromic. These sequences can be used to isolate
the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
protein associated with T-cell activation in mammalian T-cells, and is
also immunologically cross reactive with the heat shock protein Hsc70.
p74 is found in a limited number of cell types, but is particularly
expressed on B and T cells. p74 can be isolated by lysis of a suitable
cell with an amphoteric detergent, and then passed through an affinity
column containing a covalently bound HLA-B2702 palindromic peptide.
Compositions comprising the extracellular fragment of p74 combined with
HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits
cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate
compounds can be screened for their effect on the cytolytic activity of
T-cells, by combining them with the extracellular portion of p74 and
determining the amount of binding between the candidate compound and p74.
Modulation of CTL activity can be inhibited in a cellular composition
containing T-cells and antigen presenting cells (APCs), by adding to the
mix the extracellular portion of p74, in an amount sufficient to compete
with p74 for the binding of the p74 ligand.
SQ Sequence 12 AA;

Query Match 100.0%; Score 29; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
| | | | |
DB 7 RIALRY 12

RESULT 15
W33798
ID W33798 standard; peptide; 12 AA.
AC W33798;
DT 19-JUN-1998 (first entry)
DE Peptide B2702.84-79/79-84 tested for immunomodulating activity.
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
KW rejection.
OS Synthetic.
OS Homo sapiens.
PN WO9744351-A1.
PD 27-NOV-1997.
PF 22-MAY-1997; U08689.
PR 24-MAY-1996; US-653294.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Beulow R, Clayberger C, Krensky AM;
DR WPI: 98-086530/08.
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
alpha-1 domain, used for preventing rejection of transplants or
treating autoimmune diseases
PT Example 1; Page 19; 41pp; English.
PS Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
activity. A peptide-type compound or variant is claimed which has
immunomodulating activity, including the N-terminal acylated and/or
C-terminal amidated or esterified forms of up to 60 amino acids, where
the peptide-type compound comprises the formula: A-B, where A, B =
(R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or

CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
acid. The sequence in the brackets may optionally be absent or truncated
at any peptide type bond within the brackets. The compounds comprise
amino acid sequences related to a Class I HLA-B alpha domain (positions
79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
undesirably attacking cells in a host or in vitro. They can also be
used in combination with antigenic peptides or proteins of interest to
activate CTLs. They can also inhibit the proliferation of T cells in
response to anti-CD3. The peptide can be used for preventing rejection
of transplants or for treating autoimmune diseases, e.g. diabetes,
rheumatoid arthritis and lupus erythematosus. The products can also be
used for detection and diagnosis.
SQ Sequence 12 AA;

Query Match 100.0%; Score 29; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
| | | | |
DB 7 RIALRY 12

Search completed: February 8, 2000, 01:29:35
Job time: 1747 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 11:54:13 ; Search time 117.7 Seconds
(without alignments)
2.405 Million cell updates/sec

Title: US-08-653-294-4

Perfect score: 29

Sequence: 1 RIALRY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

PIR_62:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 29 | 100.0 | 237 | 2 A71082 | hypothetical prote |
| 2 | 29 | 100.0 | 273 | 2 I38509 | MHC class I histoc |
| 3 | 29 | 100.0 | 274 | 1 HLH032 | MHC class I histoc |
| 4 | 29 | 100.0 | 274 | 2 I54463 | MHC HLA-B*38 chain |
| 5 | 29 | 100.0 | 354 | 2 I59308 | class I histocompa |
| 6 | 29 | 100.0 | 354 | 2 I80168 | class I histocompa |
| 7 | 29 | 100.0 | 354 | 2 I80167 | class I histocompa |
| 8 | 29 | 100.0 | 355 | 2 I80169 | class I histocompa |
| 9 | 29 | 100.0 | 355 | 2 I80171 | class I histocompa |
| 10 | 29 | 100.0 | 359 | 1 HLH012 | MHC class I histoc |
| 11 | 29 | 100.0 | 362 | 1 HLH088 | MHC class I histoc |
| 12 | 29 | 100.0 | 362 | 2 B30345 | MHC class I histoc |
| 13 | 29 | 100.0 | 362 | 2 JH0541 | class I histocompa |
| 14 | 29 | 100.0 | 362 | 2 JH0539 | class I histocompa |
| 15 | 29 | 100.0 | 362 | 2 JH0540 | class I histocompa |
| 16 | 29 | 100.0 | 362 | 2 A45834 | MHC class I histoc |
| 17 | 29 | 100.0 | 362 | 2 I84486 | transmembrane glyc |
| 18 | 29 | 100.0 | 362 | 2 I62045 | gene HLA B-1517 pr |
| 19 | 29 | 100.0 | 362 | 2 I84490 | lymphocyte antigen |
| 20 | 29 | 100.0 | 362 | 2 I37521 | HLA-B*57.2 antigen |
| 21 | 29 | 100.0 | 362 | 2 A30345 | MHC class I histoc |
| 22 | 29 | 100.0 | 362 | 2 I59633 | MHC HLA-B transmem |
| 23 | 29 | 100.0 | 362 | 2 S24434 | class I histocompa |
| 24 | 29 | 100.0 | 362 | 2 I37120 | MHC class I histoc |
| 25 | 29 | 100.0 | 363 | 2 S07113 | class I histocompa |
| 26 | 29 | 100.0 | 363 | 2 S03537 | class I histocompa |
| 27 | 29 | 100.0 | 364 | 2 A35997 | MHC class I histoc |
| 28 | 29 | 100.0 | 364 | 2 D35997 | MHC class I histoc |
| 29 | 29 | 100.0 | 365 | 2 S77963 | MHC class I histoc |
| 30 | 29 | 100.0 | 365 | 2 JH0537 | class I histocompa |

31 29 100.0 365 2 I54416
32 29 100.0 365 2 I54493
33 28 96.6 337 2 JC5832
34 27 93.1 85 1 W98PB7
35 27 93.1 153 2 S57579
36 27 93.1 583 2 T02209
37 27 93.1 680 2 H70347
38 27 93.1 824 2 D64738
39 26 89.7 571 2 G75165
40 26 89.7 953 2 S13520
41 26 89.7 3573 2 S23070
42 25 86.2 79 2 S76214
43 25 86.2 111 2 B41629
44 25 86.2 111 2 A41629
45 25 86.2 131 2 I40656

ALIGNMENTS

RESULT 1

A71082
hypothetical protein PH0919 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 14-Aug-1998
C:Accession: A71082
R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Hino, Y.; Yamamoto, S.; Se
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yanazaki, J.; Kushida, N.; Ogu
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137
A:Accession: A71082
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-237 <KAW>
A:Cross-references: GB:AP000004; NID:93236131; PID:d1030958; PID:g3257332
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBa
C:Genetics:
A:Gene: PH0919

Query Match 100.0%; Score 29; DB 2; Length 237;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
|||||
DB 18 RIALRY 23

RESULT 2

I38509
MHC class I histocompatibility antigen - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 23-Jul-1999
C:Accession: I38509
R:Cereb, N.; Choi, J.W.; Riu, K.Z.; Yang, S.Y.
Tissue Antigens 44, 271-273, 1994
A:Title: HLA-B*5105, a newly identified B51 IEF variant.
A:Reference number: I38509; MUID:95176331
A:Accession: I38509
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-273 <RES>
A:Cross-references: EMBL:U06697; NID:g469544; PIDN:AAA92997.1; PID:g469545
C:Genetics:
A:Gene: GDB:HLA-B
A:Cross-references: GDB:I20048; OMIM:I42830
A:Map position: 6p21.3-6p21.3
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

```
Query Match      100.0%; Score 29; DB 2; Length 273;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
    |||||
Db 78 RIALRY 83

RESULT 3
HLH32
MHC class I histocompatibility antigen HLA-A32 alpha chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 05-Sep-1997
C:Accession: A26088
R:Wan, A.M.; Ennis, P.; Parham, P.; Holmes, N.
J. Immunol. 137, 3671-3674, 1986
A:Title: The primary structure of HLA-A32 suggests a region involved in formation of the
A:Reference number: A26088; MUID:87058961
A:Accession: A26088
A:Molecule type: protein
A:Residues: 1-274 <WAN>
C:Genetics:
A:Gene: GDB:HLA-A
A:Cross-references: GDB:119310; OMIM:142800
A:Map position: 6p21.3-6p21.3
A:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: glycoprotein; heterodimer; transmembrane protein; transplantation antigen
F:196-261/Domain: immunoglobulin homology <IMM>
F:86/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      100.0%; Score 29; DB 1; Length 274;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
    |||||
Db 79 RIALRY 84

RESULT 4
I54463
MHC HLA-B38 chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
C:Accession: I54463
R:Mueller, C.A.; Engler-Blum, G.; Gekeler, V.; Steiert, I.; Weiss, E.; Schmidt, H.
Immunogenetics 30, 200-207, 1989
A:Title: Genetic and serological heterogeneity of the supertypic HLA-B locus specificities
A:Reference number: I54463; MUID:89379286
A:Accession: I54463
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-274 <RES>
A:Cross-references: GB:M29864; NID:g187674; PIDN:AAA36222.1; PID:g187675
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match      100.0%; Score 29; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
    |||||
Db 79 RIALRY 84

RESULT 5
I59308
Class I histocompatibility antigen - pygmy chimpanzee (fragment)
C:Species: Pan paniscus (pygmy chimpanzee, bonobo)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 23-Jul-1999
```

```
C:Accession: I59308
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Wat
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I59308
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-354 <RES>
A:Cross-references: EMBL:U05575; NID:g454767; PIDN:AAA50178.1; PID:g454768
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match      100.0%; Score 29; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
    |||||
Db 95 RIALRY 100

RESULT 6
I80168
Class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999
C:Accession: I80168
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Wat
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I80168
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-354 <RES>
A:Cross-references: EMBL:U05579; NID:g454775; PIDN:AAA50182.1; PID:g454776
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
```

```
Query Match      100.0%; Score 29; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
    |||||
Db 95 RIALRY 100

RESULT 7
I80167
Class I histocompatibility antigen - pygmy chimpanzee (fragment)
C:Species: Pan paniscus (pygmy chimpanzee, bonobo)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999
C:Accession: I80167
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Wat
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I80167
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-354 <RES>
A:Cross-references: EMBL:U05578; NID:g454773; PIDN:AAA50181.1; PID:g454774
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
```

```
Query Match      100.0%; Score 29; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
    |||||
```

Db 95 RIALRY 100

RESULT 8

180169
class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999
C:Accession: I80169
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkins
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I80169
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-355 <RES>
A:Cross-references: EMBL:U05580; NID:g454777; PIDN:AAA50183.1; PID:g454778
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 29; DB 2; Length 355;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6

Db 95 RIALRY 100

RESULT 9

180171
class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999
C:Accession: I80171
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkins
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I80171
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-355 <RES>
A:Cross-references: EMBL:U05582; NID:g454781; PIDN:AAA50185.1; PID:g454782
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 29; DB 2; Length 355;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6

Db 95 RIALRY 100

RESULT 10

HLH012
MHC class I histocompatibility antigen HLA alpha chain precursor (clone pHLA 12.4) - human
C:Species: Homo sapiens (man)
C:Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 22-Jun-1999
C:Accession: A02189
R:Malissen, M.; Malissen, B.; Jordan, B.R.
Proc. Natl. Acad. Sci. U.S.A. 79, 893-897, 1982
A:Title: Exon/intron organization and complete nucleotide sequence of an HLA gene.
A:Reference number: A02189; MUID:82151002
A:Accession: A02189
A:Molecule type: DNA
A:Residues: 1-359 <MAL>
A:Cross-references: GB:J00191; GB:V00526; NID:g187600; PIDN:AAA36218.1; PID:g386873
C:Comment: The seven exons correspond approximately to the domain structure of this chain
C:Genetics:

A:Map position: 6p21.3

A:Introns: 22/1; 112/1; 204/1; 296/1; 335/1; 346/1
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplant
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-359/Product: class I histocompatibility antigen HLA alpha chain #status predicted
F:22-304/Domain: extracellular #status predicted <EXT>
F:22-111/Domain: alpha-1 <EX1>
F:112-203/Domain: alpha-2 <EX2>
F:217-282/Domain: immunoglobulin homology <IMM>
F:305-329/Domain: transmembrane #status predicted <TM>
F:335-359/Domain: intracellular #status predicted <INT>
F:107/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:224-280/Disulfide bonds: #status predicted

Query Match 100.0%; Score 29; DB 1; Length 359;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6

Db 100 RIALRY 105

RESULT 11

HLH088
MHC class I histocompatibility antigen HLA-Bw58 alpha chain precursor - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 05-Sep-1997
C:Accession: A23895
R:Ways, J.P.; Coppin, H.L.; Parham, P.
J. Biol. Chem. 260, 11924-11933, 1985
A:Title: The complete primary structure of HLA-Bw58.
A:Reference number: A23895; MUID:86008247
A:Accession: A23895
A:Molecule type: DNA
A:Residues: 1-362 <WAY>
A:Note: the authors translated the codon GGC for residue 349 as Ser
C:Comment: This protein is a subtype of the HLA-B17 family.
C:Genetics:

A:Gene: GDB:HLA-B
A:Cross-references: GDB:120048; OMIM:142830
A:Map position: 6p21.3-6p21.3

A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplant
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-362/Product: class I histocompatibility antigen HLA-Bw58 alpha chain #status pre
F:25-307/Domain: extracellular #status predicted <EXT>
F:25-114/Domain: alpha-1 <EX1>
F:115-206/Domain: alpha-2 <EX2>
F:220-285/Domain: immunoglobulin homology <IMM>
F:308-331/Domain: transmembrane #status predicted <TM>
F:332-362/Domain: intracellular #status predicted <INT>
F:110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 29; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6

Db 103 RIALRY 108

RESULT 12

B30345
MHC class I histocompatibility antigen HLA-Bw52 precursor - human
C:Species: Homo sapiens (man)
C:Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 16-Feb-1997
C:Accession: B30345

R:Hayashi, H.; Ennis, P.D.; Ariga, H.; Salter, R.D.; Parham, P.; Kano, K.; Takiguchi, M.
 J. Immunol. 142, 306-311, 1989
 A:Title: HLA-B*51 and HLA-B*52 differ by only two amino acids which are in the helical re
 A:Reference number: A30345; MUID:89080265

A:Accession: B30345
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-362 <HAY>
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C:Keywords: transmembrane protein
 F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 29; DB 2; Length 362;
 Best Local Similarity 100.0%; Pred. No. 7.5;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 |||||
 Db 103 RIALRY 108

RESULT 13

JH0541
 Class I histocompatibility antigen Gogo-B0103 heavy chain precursor - lowland gorilla
 C:Species: Gorilla gorilla gorilla (lowland gorilla)
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
 C:Accession: JH0541
 R:Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.
 J. Exp. Med. 174, 1491-1509, 1991
 A:Title: Gorilla class I major histocompatibility complex alleles: comparison to human a
 A:Reference number: JH0534; MUID:92078860
 A:Accession: JH0541

A:Molecule type: DNA
 A:Residues: 1-362 <LAW>

A:Cross-references: EMBL:X60254; NID:g22869; PIDN:CAA42806.1; PID:g22870
 A:Experimental source: EBV-transformed B cell
 C:Genetics:

A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C:Keywords: transmembrane protein
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0103 #status pre
 F:25-114/Domain: alpha-1 <AL1>
 F:115-206/Domain: alpha-2 <AL2>
 F:207-298/Domain: alpha-3 <AL3>
 F:220-285/Domain: immunoglobulin homology <IMM>
 F:299-362/Domain: intracellular #status predicted <INT>

Query Match 100.0%; Score 29; DB 2; Length 362;
 Best Local Similarity 100.0%; Pred. No. 7.5;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 |||||
 Db 103 RIALRY 108

RESULT 14

JH0539
 Class I histocompatibility antigen Gogo-B0101 heavy chain precursor - lowland gorilla
 C:Species: Gorilla gorilla gorilla (lowland gorilla)
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
 C:Accession: JH0539
 R:Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.
 J. Exp. Med. 174, 1491-1509, 1991
 A:Title: Gorilla class I major histocompatibility complex alleles: comparison to human a
 A:Reference number: JH0534; MUID:92078860
 A:Accession: JH0539

A:Molecule type: DNA
 A:Residues: 1-362 <LAW>

A:Cross-references: EMBL:X60255; NID:g22865; PIDN:CAA42807.1; PID:g22866

A:Experimental source: EBV-transformed B cell
 C:Genetics:
 A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C:Keywords: transmembrane protein
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0101 #status
 F:25-114/Domain: alpha-1 <AL1>
 F:115-206/Domain: alpha-2 <AL2>
 F:207-298/Domain: alpha-3 <AL3>
 F:220-285/Domain: immunoglobulin homology <IMM>
 F:299-362/Domain: intracellular #status predicted <INT>

Query Match 100.0%; Score 29; DB 2; Length 362;
 Best Local Similarity 100.0%; Pred. No. 7.5;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 |||||
 Db 103 RIALRY 108

RESULT 15

JH0540
 Class I histocompatibility antigen Gogo-B0102 heavy chain precursor - lowland gorilla
 C:Species: Gorilla gorilla gorilla (lowland gorilla)
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
 C:Accession: JH0540
 R:Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.
 J. Exp. Med. 174, 1491-1509, 1991
 A:Title: Gorilla class I major histocompatibility complex alleles: comparison to huma
 A:Reference number: JH0534; MUID:92078860
 A:Accession: JH0540

A:Molecule type: DNA
 A:Residues: 1-362 <LAW>

A:Cross-references: EMBL:X60693; NID:g22867; PIDN:CAA43101.1; PID:g22868
 A:Experimental source: EBV-transformed B cell
 C:Genetics:

A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C:Keywords: transmembrane protein
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0102 #status
 F:25-114/Domain: alpha-1 <AL1>
 F:115-206/Domain: alpha-2 <AL2>
 F:207-298/Domain: alpha-3 <AL3>
 F:220-285/Domain: immunoglobulin homology <IMM>
 F:299-362/Domain: intracellular #status predicted <INT>

Query Match 100.0%; Score 29; DB 2; Length 362;
 Best Local Similarity 100.0%; Pred. No. 7.5;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 |||||
 Db 103 RIALRY 108

Search completed: February 7, 2000, 11:54:14
 Job time: 24324 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 00:59:44 ; Search time 63.71 Seconds
(without alignments)
2.813 Million cell updates/sec

Title: US-08-653-294-4

Perfect score: 29
Sequence: 1 RIALRY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_38.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 29 | 100.0 | 359 | 1 B01_PANTR | P13750 pan troglod |
| 2 | 29 | 100.0 | 362 | 1 B01_GORGO | P30379 gorilla gor |
| 3 | 29 | 100.0 | 362 | 1 B02_GORGO | P30380 gorilla gor |
| 4 | 29 | 100.0 | 362 | 1 B03_GORGO | P10317 homo sapien |
| 5 | 29 | 100.0 | 362 | 1 B15_HUMAN | P30487 homo sapien |
| 6 | 29 | 100.0 | 362 | 1 B47_HUMAN | P18464 homo sapien |
| 7 | 29 | 100.0 | 362 | 1 B49_HUMAN | P30489 homo sapien |
| 8 | 29 | 100.0 | 362 | 1 B52_HUMAN | P30490 homo sapien |
| 9 | 29 | 100.0 | 362 | 1 B53_HUMAN | P30491 homo sapien |
| 10 | 29 | 100.0 | 362 | 1 B54_HUMAN | P18465 homo sapien |
| 11 | 29 | 100.0 | 362 | 1 B60_HUMAN | P30497 homo sapien |
| 12 | 29 | 100.0 | 362 | 1 B61_HUMAN | P10319 homo sapien |
| 13 | 29 | 100.0 | 362 | 1 B62_HUMAN | P01893 homo sapien |
| 14 | 29 | 100.0 | 362 | 1 HLAH_HUMAN | P30378 gorilla gor |
| 15 | 29 | 100.0 | 365 | 1 IA04_GORGO | P30447 homo sapien |
| 16 | 29 | 100.0 | 365 | 1 IA23_HUMAN | P05534 homo sapien |
| 17 | 29 | 100.0 | 365 | 1 IA24_HUMAN | P18462 homo sapien |
| 18 | 29 | 100.0 | 365 | 1 IA25_HUMAN | P10314 homo sapien |
| 19 | 29 | 100.0 | 365 | 1 IA32_HUMAN | P03789 bacterioph |
| 20 | 27 | 93.1 | 85 | 1 V192_BPT7 | P37024 escherichia |
| 21 | 27 | 93.1 | 809 | 1 HRPB_ECOLI | P23514 rattus norv |
| 22 | 26 | 89.7 | 953 | 1 COPB_RAT | Q03132 saccharopol |
| 23 | 26 | 89.7 | 3567 | 1 ERY2_SACER | Q03132 saccharopol |
| 24 | 25 | 86.2 | 268 | 1 YRY3_CAEEL | O10007 caenorhabdi |
| 25 | 25 | 86.2 | 322 | 1 M1AA_PSEPU | O30762 pseudomonas |
| 26 | 25 | 86.2 | 331 | 1 UL38_HCMVA | P16779 human cytom |
| 27 | 25 | 86.2 | 349 | 1 LACH_SCHAM | Q28474 schistosom |
| 28 | 25 | 86.2 | 462 | 1 COXX_YEAST | P21592 saccharomyc |
| 29 | 25 | 86.2 | 638 | 1 G0IM_COXBU | P45650 coxiella bu |
| 30 | 25 | 86.2 | 959 | 1 MSH1_YEAST | P25846 saccharomyc |
| 31 | 25 | 86.2 | 967 | 1 HPA_ECOLI | P23852 escherichia |
| 32 | 25 | 86.2 | 974 | 1 HPA_ECOLI | O43306 homo sapien |
| 33 | 25 | 86.2 | 1165 | 1 CYA6_CANFA | P30804 canis famil |
| 34 | 25 | 86.2 | 1165 | 1 CYA6_MOUSE | Q01341 mus musculus |

ALIGNMENTS

| RESULT | 1 | 1166 | 1166 | 1 | CYA6_RAT | Q03343 | rattus norv |
|---|----|------|------|------|--------------|--------|--------------|
| 1B01_PANTR | 35 | 25 | 86.2 | 2225 | 1 PYR1_HUMAN | P27708 | homo sapien |
| ID 1B01_PANTR | 36 | 25 | 86.2 | 2225 | 1 PYR1_MESAU | P08955 | mesocricetu |
| AC P13750; | 37 | 25 | 86.2 | 2569 | 1 LMA3_MOUSE | Q61789 | mus musculus |
| DT 01-JAN-1990 (Rel. 13, Created) | 38 | 25 | 86.2 | 3712 | 1 LMA_DROME | Q00174 | drosophila |
| DT 01-JAN-1990 (Rel. 13, Last sequence update) | 39 | 25 | 86.2 | 108 | 1 RS25_YEAST | P07282 | saccharomyc |
| DT 01-APR-1993 (Rel. 25, Last annotation update) | 40 | 24 | 82.8 | 156 | 1 RM25_YEAST | P23369 | saccharomyc |
| DE CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-1 ALPHA CHAIN PRECURSOR | 41 | 24 | 82.8 | 213 | 1 COBO_PSEDE | P29930 | pseudomonas |
| DE (FRAGMENT) | 42 | 24 | 82.8 | 274 | 1 URED_PROMI | P17089 | proteus mir |
| OS Pan troglodytes (Chimpanzee) | 43 | 24 | 82.8 | 288 | 1 YGHF_ECOLI | Q46834 | escherichia |
| OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; | 44 | 24 | 82.8 | 298 | 1 HAY1_MOUSE | P01895 | mus musculus |
| OC Eutheria; Primates; Catarrhini; Homiidae; Pan. | 45 | 24 | 82.8 | | | | |
| RN [1] | | | | | | | |
| RP SEQUENCE FROM N.A. | | | | | | | |
| RX MEDLINE; 89030641. | | | | | | | |
| RA MAYER W.E., JONKER M., KLEIN D., IVANYI P., VAN SEVENTER G., | | | | | | | |
| RA KLEIN J.; | | | | | | | |
| RT "Nucleotide sequences of chimpanzee MHC class I alleles: evidence for | | | | | | | |
| RT trans-species mode of evolution."; | | | | | | | |
| RL EMBO J. 7:2765-2774(1988). | | | | | | | |
| RN [2] | | | | | | | |
| RP REVISIONS. | | | | | | | |
| RA MAYER W.; | | | | | | | |
| RL Submitted (FEB-1989) to the EMBL/GenBank/DBJ databases. | | | | | | | |
| CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO | | | | | | | |
| CC THE IMMUNE SYSTEM. | | | | | | | |
| CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2- | | | | | | | |
| CC MICROGLOBULIN). | | | | | | | |
| CC This SWISS-PROT entry is copyright. It is produced through a collaboration | | | | | | | |
| CC between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | | | | | |
| CC the European Bioinformatics Institute. There are no restrictions on its | | | | | | | |
| CC use by non-profit institutions as long as its content is in no way | | | | | | | |
| CC modified and this statement is not removed. Usage by and for commercial | | | | | | | |
| CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | | | | | |
| CC or send an email to license@isb-sib.ch) | | | | | | | |
| CC ----- | | | | | | | |
| CC EMBL; X13115; CAA31507.1; - | | | | | | | |
| DR PIR; S03537; S03537. | | | | | | | |
| DR HSSP; P03989; 1HSA. | | | | | | | |
| DR PROSITE; PS00290; IG_MHC; 1. | | | | | | | |
| DR PFAM; PF00047; ig; 1. | | | | | | | |
| DR PFAM; PF00129; MHC_1; 1. | | | | | | | |
| KW MHC I; Transmembrane; Glycoprotein; Signal. | | | | | | | |
| FT NON_TER | | | | | | | |
| FT SIGNAL | | | | | | | |
| CHAIN | | | | | | | |
| FT DOMAIN | | | | | | | |
| FT DOMAIN | | | | | | | |
| FT DOMAIN | | | | | | | |
| FT DOMAIN | | | | | | | |
| FT TRANSMEM | | | | | | | |
| FT DOMAIN | | | | | | | |
| FT DISULFID | | | | | | | |
| FT DISULFID | | | | | | | |
| FT CARBOHYD | | | | | | | |

CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
B-1 ALPHA CHAIN.
EXTRACELLULAR ALPHA-1.
EXTRACELLULAR ALPHA-2.
EXTRACELLULAR ALPHA-3.
CONNECTING PEPTIDE.
CYTOPLASMIC TAIL.
BY SIMILARITY.
BY SIMILARITY.

SQ SEQUENCE 359 AA; 40173 MW; 5395FFC9 CRC32;

Query Match 100.0%; Score 29; DB 1; Length 359;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 6; Conservative 0; Mismatches 0; Indels

| | |
|----|---------------|
| Qy | 1 RIALRY 6 |
| | |
| Db | 99 RIALRY 104 |

| | |
|------------|---|
| RESULT | 2 |
| IBO1_GORGO | |
| ID | IBO1_GORGO |
| AC | P30379; |
| DT | 01-APR-1993 (Rel. 25, Created) |
| DT | 01-APR-1993 (Rel. 25, Last sequence update) |
| DT | 01-APR-1993 (Rel. 25, Last annotation update) |
| DE | CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0101 ALPHA CHAIN PRECURSOR. |
| OS | Gorilla gorilla gorilla (Lowland gorilla). |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; |
| NC | Eutheria; Primates; Catarrhini; Hominidae; Gorilla. |
| RN | [1] |

```

RD SEQUENCE FROM N.A.
RX MEDLINE; 92078660.
RA LAWLOW D.A., WARREN E., TAYLOR P., PARHAM P.;
RT "Gorilla class I major histocompatibility complex alleles: comparison
RT to human and chimpanzee class I.";
RL J. Exp. Med. 174:1491-1509(1991).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).

```

| | | | |
|----|---|---------|---------------------------|
| CC | EMBL; X60255; CAA42807.1; -. | | |
| DR | PIR; JH0539; JH0539. | | |
| DR | HSSP; P03989; IHSA. | | |
| DR | PROSITE; PS00290; IG_MHC; 1. | | |
| DR | PFAM; PF00047; ig; 1. | | |
| DR | PFAM; PF00129; MHC_1; 1. | | |
| DR | MHC I; Transmembrane; Glycoprotein; Signal. | | |
| FW | SIGNAL | 1 | 24 |
| FT | | | |
| FT | CHAIN | 25 | 362 |
| FT | | | |
| FT | DOMAIN | 25 | 114 |
| FT | DOMAIN | 115 | 206 |
| FT | DOMAIN | 207 | 298 |
| FT | DOMAIN | 299 | 308 |
| FT | TRANSMEM | 309 | 332 |
| FT | DOMAIN | 333 | 362 |
| FT | DISULFID | 125 | 188 |
| FT | DISULFID | 227 | 283 |
| FT | CARBOHYD | 110 | 110 |
| SQ | SEQUENCE | 362 AA; | 40170 MW; 2E33E2B8 CRC32; |

```
Query Match      100.0%; Score 29; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

| | |
|----------------|----------------|
| Qy | 1 RIALRY 6 |
| | |
| D _b | 103 RIALRY 108 |

Db 103 RIALRY 108

RESULT 3

| | | | |
|------------|---|------|---------|
| 1B02_GORGO | STANDARD; | PRT; | 362 AA. |
| ID | 1B02_GORGO | | |
| AC | P30380; | | |
| DT | 01-APR-1993 (Rel. 25, Created) | | |
| DT | 01-APR-1993 (Rel. 25, Last sequence update) | | |
| DT | 01-APR-1993 (Rel. 25, Last annotation update) | | |
| DE | CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0102 ALPHA CHAIN PRECURSOR. | | |
| OS | Gorilla gorilla gorilla (Lowland gorilla). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; | | |
| OC | Eutheria; Primates; Catarrhini; Homiidae; Gorilla. | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RX | MEDLINE; 92078860. | | |
| RA | LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.; | | |
| RT | "Gorilla class I major histocompatibility complex alleles: comparison | | |
| RT | to human and chimpanzee class I."; | | |
| RL | J. Exp. Med. 174:1491-1509(1991). | | |
| CC | -I- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO | | |
| CC | THE IMMUNE SYSTEM. | | |
| CC | -I- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2- | | |
| CC | MICROGLOBULIN). | | |

```

CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; X60693; CAA43101.1; -
CC      PIR; JH0540; JH0540.
CC      HSP; P03989; 1HSA.
CC      PROSITE; PS00290; IG_MHC; 1.
CC      PFAM; PF00047; Ig; 1.
CC      PFAM; PF00129; MHC_I; 1.
CC      MHC I; Transmembrane; Glycoprotein; Signal.
CC      SIGNAL 1 24 BY SIMILARITY.
CC      CHAIN 25 362 CLASS I HISTOCOMPATIBILITY ANTIGEN,
CC
CC      DOMAIN 25 114 GOGO-B0102 ALPHA CHAIN.
CC      DOMAIN 115 206 EXTRACELLULAR ALPHA-1.
CC      DOMAIN 207 298 EXTRACELLULAR ALPHA-2.
CC      DOMAIN 299 308 EXTRACELLULAR ALPHA-3.
CC      TRANSMEM 309 332 CONNECTING PEPTIDE.
CC      DOMAIN 333 362 CYTOPLASMIC TAIL.
CC      DISULFID 125 188 BY SIMILARITY.
CC      DISULFID 227 283 BY SIMILARITY.
CC      CARBOHYD 110 110 BY SIMILARITY.
CC      SEQUENCE 362 AA; 40204 MW; 3CF119AD CRC32;
CC

```

| | | | | |
|-----------------------|--------------|--------------|------------|-------------|
| Query Match | 100.0%; | Score 29; | DB 1; | Length 362; |
| Best Local Similarity | 100.0%; | pred. NO. 4; | | |
| Matches 6: | Conservative | 0; | Mismatches | 0; |
| | | | Indels | 0; |
| | | | Gaps | 0; |

| | |
|----|----------------|
| QY | 1 RIALRY 6 |
| | |
| Db | 103 RIALRY 108 |

RESULT

| | | | |
|------------|---|------|---------|
| 1B03_GORGO | STANDARD; | PRT; | 362 AA. |
| ID | 1B03_GORGO | | |
| AC | F30361; | | |
| DT | 01-APR-1993 (Rel. 25, Created) | | |
| DT | 01-APR-1993 (Rel. 25, Last sequence update) | | |
| DT | 01-APR-1993 (Rel. 25, Last annotation update) | | |
| DE | CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0103 | | |
| DE | CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0103 | | |
| DE | ALPHA CHAIN PRECURSOR. | | |
| OS | Gorilla gorilla gorilla (Lowland gorilla). | | |
| OS | Gorilla gorilla gorilla (Lowland gorilla). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; | | |

OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92078660.
 RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;
 RT "Gorilla class I major histocompatibility complex alleles: comparison
 to human and chimpanzee class I";
 RL J. Exp. Med. 174:1491-1509(1991).
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X60254; CAA42806.1; -;
 DR FIR; JH0541; JH0541.
 DR HSP; P03989; IHSA.
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; IG_1;
 DR PFAM; PF00129; MHC_1; 1.
 KW MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT -----
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 238
 FT DOMAIN 299 308
 FT TRANSMEM 309 332
 FT DOMAIN 333 362
 FT DISULFID 125 188
 FT DISULFID 227 283
 FT CARBOHYD 110 110
 SQ SEQUENCE 362 AA; 40248 MW; FEA6A941 CRC32;

 Query Match 100.0%; Score 29; DB 1; Length 362;
 Best Local Similarity 100.0%; Pred. No. 4;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 RIALRY 6
 Db 103 RIALRY 108

 RESULT 5
 ID 1B15_HUMAN
 AC P10317;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-27 B*2702 ALPHA CHAIN
 DE PRECURSOR (B-27K) (B27.2).
 GN HLA-B OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86220133.
 RA SEEMANN G.H.A., REIN R.S., BROWN C.S., PLOEGH H.L.;
 RT "Gene conversion-like mechanisms may generate polymorphism in human
 class I genes";
 RL EMBO J. 5:547-552(1986).
 RN [2]

RP SEQUENCE FROM N.A.
 RA PARHAM P., ARNETT K.L., ADAMS E.J.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 86-107 AND 171-181.
 RX MEDLINE; 86042671.
 RA VEGA M.A., EZQUERRA A., ROJO S., APARICIO P., BRAGADO R.,
 RA LOPEZ DE CASTRO J.A.;
 RT "Structural analysis of an HLA-B27 functional variant: identification
 of residues that contribute to the specificity of recognition by
 cytolytic T lymphocytes";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7394-7398(1985).
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X03664; CAA27301.1; -;
 DR EMBL; X03667; CAA27301.1; JOINED.
 DR EMBL; L38504; AAA69724.1; -;
 DR FIR; B25092; HLHUBK.
 DR HSP; P03989; IHSA.
 DR MIM; 142830;
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; IG_1;
 DR PFAM; PF00129; MHC_1; 1.
 KW MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT -----
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT TRANSMEM 309 332
 FT DOMAIN 333 362
 FT CARBOHYD 110 110
 FT DISULFID 125 188
 FT DISULFID 227 283
 SQ SEQUENCE 362 AA; 40397 MW; 9798F0BB CRC32;

 Query Match 100.0%; Score 29; DB 1; Length 362;
 Best Local Similarity 100.0%; Pred. No. 4;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 RIALRY 6
 Db 103 RIALRY 108

 RESULT 6
 ID 1B47_HUMAN
 AC P30487;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-49(B-21) B*4901 ALPHA CHAIN
 DE PRECURSOR
 GN HLA-B OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]


```

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-52(B-5) B*5201 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 89080265.
RA HAYASHI H., ENNIS P.D., ARIGA H., SALTER R.D., PARHAM P., KANO K.,
RA TAKIGUCHI M.;
RT "HLA-B51 and HLA-BW52 differ by only two amino acids which are in the
RT helical region of the alpha 1 domain.";
RL J. Immunol. 142:306-311(1989).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22799; AAA59645.1; ALT SEQ.
DR EMBL; M22793; AAA59645.1; JOINED.
DR EMBL; M22794; AAA59645.1; JOINED.
DR EMBL; M22795; AAA59645.1; JOINED.
DR EMBL; M22796; AAA59645.1; JOINED.
DR EMBL; M22797; AAA59645.1; JOINED.
DR EMBL; M22798; AAA59645.1; JOINED.
DR PIR; B30345; B30345.
DR PIR; B30548; B30548.
DR HSSP; P30491; 1A1M.
DR MIN; 142830; -.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
DR MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362
FT FT
FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT B-51(B-5) B*5104 ALPHA CHAIN.
FT EXTRACELLULAR ALPHA-1.
FT DOMAIN 25 114
FT DOMAIN 115 206
FT DOMAIN 207 298
FT DOMAIN 299 308
FT TRANSMEM 309 332
FT DOMAIN 333 362
FT CARBOHYD 110 110
FT BY SIMILARITY.
FT DISULFID 125 188
FT DISULFID 227 283
FT BY SIMILARITY.
SQ SEQUENCE 362 AA; 40521 MW; 3B436FE8 CRC32;

Query Match 100.0%; Score 29; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
DB 103 RIALRY 108
-----
RESULT 10
ID 1B54_HUMAN STANDARD; PRT; 362 AA.
AC P30491;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-51(B-5) B*5104 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92269955.
RA BELICH M.P., MADRIGAL J.A., HILDEBRAND W.H., ZEMMOUR J.,
RA WILLIAMS R.C., LUZ R., PETZL-ERLER M.L., PARHAM P.;
RT "Unusual HLA-B alleles in two tribes of Brazilian Indians.";
RL Nature 357:326-329(1992).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z15143; CAA78849.1; -.
DR HSSP; P30491; 1A1M.
DR MIN; 142830; -.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
DR MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362
FT FT
FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT B-51(B-5) B*5104 ALPHA CHAIN.
FT EXTRACELLULAR ALPHA-1.
FT DOMAIN 25 114
FT DOMAIN 115 206
FT DOMAIN 207 298
FT DOMAIN 299 308
FT TRANSMEM 309 332
FT DOMAIN 333 362
FT CARBOHYD 110 110
FT BY SIMILARITY.
FT DISULFID 125 188
FT DISULFID 227 283
FT BY SIMILARITY.
SQ SEQUENCE 362 AA; 40560 MW; F22F09AB CRC32;

Query Match 100.0%; Score 29; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
DB 103 RIALRY 108
-----
RESULT 9
ID 1B53_HUMAN STANDARD; PRT; 362 AA.
AC P30490;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)

```

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-53 B*5301 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91033941.
RA HAYASHI H., Ooba T., NAKAYAMA S., SEKIMATA M., KANO K.,
RA TAKIGUCHI M.;
RT "Allostericities between HLA-B*53 and HLA-B*35 are generated by
RT substitution of the residues associated with HLA-B*4/B*6 public
RT epitopes.";
RL Immunogenetics 32:195-199(1990).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 25-302.
RX MEDLINE; 96209672.
RA SMITH K.J., REID S.W., HARLOS K., MCMICHAEL A.J., STUART D.I.,
RA BELL J.I., JONES E.Y.;
RT "Bound water structure and polymorphic amino acids act together to
RT allow the binding of different peptides to MHC class I HLA-B*53.";
RL Immunity 4:215-228(1996).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M58636; AAA36228.1; -
CC PIR; A45834; A45834.
CC PDB; 1A1M; 08-APR-98.
CC PDB; 1A1O; 08-APR-98.
CC MIM; 142830; -
CC PROSITE; PS00290; IG_MHC; 1.
CC PFAM; PF00047; IG; 1.
CC PFAM; PF00129; MHC_I; 1.
CC MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT BW-53 B*5301 ALPHA CHAIN.
FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
FT DOMAIN 299 308 CONNECTING PEPTIDE.
FT TRANSMEM 309 332
FT DOMAIN 333 362 CYTOPLASMIC TAIL.
FT CARBOHYD 110 110 BY SIMILARITY.
FT DISULFID 125 188
FT DISULFID 227 283
SQ SEQUENCE 362 AA; 40495 MW; 2BDC745E CRC32;

Query Match 100.0%; Score 29; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
Db 103 RIALRY 108

RESULT 11
1B60_HUMAN
ID 1B60_HUMAN STANDARD; PRT; 362 AA.
AC P18465;

DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-57(B-17) B*5701 ALPHA
DE CHAIN PRECURSOR (B*57.1).
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90207291.
RA ENNIS P.D., ZEMOUR J., SALTER R.D., PARHAM P.;
RT "Rapid cloning of HLA-A,B CDNA by using the polymerase chain
RT reaction: frequency and nature of errors produced in amplification.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91067476.
RA ISAWAT M., GIRDLESTONE J., MILSTEIN C.;
RT "Nucleotide sequence of an HLA-B*57 gene.";
RL Nucleic Acids Res. 18:6702-6702(1990).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M32318; AAA36231.1; -
CC EMBL; X55711; CAA39244.1; -
CC PIR; S12622; S12622.
CC PIR; D35997; D35997.
CC HSP; P30491; IALN.
CC MIM; 142830; -
CC PROSITE; PS00290; IG_MHC; 1.
CC PFAM; PF00047; IG; 1.
CC PFAM; PF00129; MHC_I; 1.
CC MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT B-57(B-17) B*5701 ALPHA CHAIN.
FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
FT DOMAIN 299 308 CONNECTING PEPTIDE.
FT TRANSMEM 309 332
FT DOMAIN 333 362 CYTOPLASMIC TAIL.
FT CARBOHYD 110 110 BY SIMILARITY.
FT DISULFID 125 188 BY SIMILARITY.
FT DISULFID 227 283 BY SIMILARITY.
SQ SEQUENCE 362 AA; 40224 MW; D91DF8DD CRC32;

Query Match 100.0%; Score 29; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
Db 103 RIALRY 108

RESULT 12
1B61_HUMAN
ID 1B61_HUMAN STANDARD; PRT; 362 AA.
AC P30497;

DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-57(B-17) B*5702 ALPHA CHAIN
 DE PRECURSOR (BW57.2).
 GN HLA-B OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 93036508.
 RA MADRIGAL J.A., ZELICH M.P., HILDEBRAND W.H., BENJAMIN R.J.,
 RA LITTLE A.M., BEMOUR J., ENNIS P.D., WARD F.E., PETZL-ERLER M.L.,
 RA MARTELL R.W., DU TOIT E.D., PARHAM P.;
 RT "Distinctive HLA-A-B antigens of black populations formed by
 RT interallelic conversion.";
 RL J. Immunol. 149:3411-3415(1992).
 CC
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X61707; CAA43876.1; -.
 DR PIR; S16774; S16774.
 DR HSP; P30491; IALM.
 DR MIN; 142830;
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; Ig; 1.
 DR MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT TRANSMEM 309 332
 FT DOMAIN 333 362
 FT CARBOHYD 110 110
 FT DISULFID 125 188
 FT DISULFID 227 283
 SQ SEQUENCE 362 AA; 40342 MW; 628C2156 CRC32;
 Query Match 100.0%; Score 29; DB 1; Length 362;
 Best Local Similarity 100.0%; Pred. No. 4;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RIALRY 6
 DB 103 RIALRY 108
 RESULT 13
 ID 1B62_HUMAN STANDARD; PRT; 362 AA.
 AC P10319;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-58(B-17) B*5801 ALPHA
 DE CHAIN PRECURSOR.
 GN HLA-B OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86008247.
 RA WAYS J.P., COPPIN H.L., PARHAM P.;
 RT "The complete primary structure of HLA-Bw58.";
 RL J. Biol. Chem. 260:11924-11933(1985).
 CC
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M11799; AAS59628.1; -.
 DR EMBL; AB008102; BAA22916.1; -.
 DR PIR; A23895; HLHUB8.
 DR HSP; P30491; IALM.
 DR MIN; 142830;
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; Ig; 1.
 DR MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT TRANSMEM 309 332
 FT DOMAIN 333 362
 FT CARBOHYD 110 110
 FT DISULFID 125 188
 FT DISULFID 227 283
 SQ SEQUENCE 362 AA; 40337 MW; 3E5E7534 CRC32;
 Query Match 100.0%; Score 29; DB 1; Length 362;
 Best Local Similarity 100.0%; Pred. No. 4;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RIALRY 6
 DB 103 RIALRY 108
 RESULT 14
 ID HLAH_HUMAN STANDARD; PRT; 362 AA.
 AC P01893;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-DEC-1999 (Rel. 39, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, ALPHA CHAIN H PRECURSOR
 DE (HLA-AR) (HLA-12.4).
 GN HLA-H OR HLAH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN SEQUENCE FROM N.A.
 RX MEDLINE: 82151002.
 RA MALISSEN M., MALISSEN B., JORDAN B.R.;
 RT "Exon/intron organization and complete nucleotide sequence of an HLA
 gene."
 RL Proc. Natl. Acad. Sci. U.S.A. 79:893-897(1982).
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 THE IMMUNE SYSTEM. COULD BE THE PRODUCT OF A PSEUDOGENE.
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 MICROGLOBULIN).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J00191; AAA36218.1; ALT_INIT.
 DR PIR: A02189; HLH012.
 DR HSSP: P03989; 1HSA.
 DR MIM: 142800; -.
 DR PROSITE: PS00290; IG_MHC; 1.
 DR PFAM: PF00047; ig; 1.
 DR PFAM: PF00129; MHC_I; 1.
 KW MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
 FT ALPHA CHAIN H.
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT DOMAIN 309 332
 FT TRANSMEM 333 362
 FT DOMAIN 333 362
 FT CYTOPLASMIC TAIL.
 FT CARBOHYD 110 110
 FT DISULFID 227 283
 FT BY SIMILARITY.
 SQ SEQUENCE 362 AA; 40850 MW; 5E610F63 CRC32;

Query Match 100.0%; Score 29; DB 1; Length 362;
 Best Local Similarity 100.0%; Pred. No. 4;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 DB 103 RIALRY 108

RESULT 15
 ID 1A04_GORGO STANDARD; PRT: 365 AA.
 AC P30378;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-A0501 ALPHA CHAIN PRECURSOR.
 OS Gorilla gorilla gorilla (Lowland gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92078860.
 RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;
 RT "Gorilla class I major histocompatibility complex alleles: comparison
 to human and chimpanzee class I."
 RL J. Exp. Med. 174:1491-1509(1991).
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 THE IMMUNE SYSTEM.
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 MICROGLOBULIN).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X60256; CAA42808.1; -.
 DR PIR: JH0537; JH0537.
 DR HSSP: P01892; 1HHH.
 DR PROSITE: PS00290; IG_MHC; 1.
 DR PFAM: PF00047; ig; 1.
 DR PFAM: PF00129; MHC_I; 1.
 KW MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 365
 FT GOGO-A0501 ALPHA CHAIN.
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT TRANSMEM 309 332
 FT DOMAIN 333 365
 FT CYTOPLASMIC TAIL.
 FT DISULFID 125 188
 FT DISULFID 227 283
 FT CARBOHYD 110 110
 FT BY SIMILARITY.
 SQ SEQUENCE 365 AA; 40895 MW; 520225DF CRC32;

Query Match 100.0%; Score 29; DB 1; Length 365;
 Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 DB 103 RIALRY 108

Search completed: February 8, 2000, 00:59:45
 Job time: 3774 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 13:17:28 ; Search time 209.03 Seconds
(without alignments)
1.990 Million cell updates/sec

Title: US-08-653-294-4

Perfect score: 29

Sequence: 1 RIALRY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL_12.*
1: sp-archaea.*
2: sp-bacteria.*
3: sp-fungi.*
4: sp-human.*
5: sp-invertebrate.*
6: sp-mammal.*
7: sp-mhc.*
8: sp-organelle.*
9: sp-phage.*
10: sp-plant.*
11: sp-todent.*
12: sp-virus.*
13: sp-vertebrate.*
14: sp-unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 29 | 100.0 | 89 | 7 | 019569 |
| 2 | 29 | 100.0 | 90 | 7 | 046697 |
| 3 | 29 | 100.0 | 133 | 7 | 019189 |
| 4 | 29 | 100.0 | 138 | 7 | 078209 |
| 5 | 29 | 100.0 | 172 | 7 | 019770 |
| 6 | 29 | 100.0 | 172 | 7 | 019774 |
| 7 | 29 | 100.0 | 172 | 7 | 019775 |
| 8 | 29 | 100.0 | 172 | 7 | 019780 |
| 9 | 29 | 100.0 | 172 | 7 | 095364 |
| 10 | 29 | 100.0 | 172 | 7 | 019771 |
| 11 | 29 | 100.0 | 172 | 7 | 019772 |
| 12 | 29 | 100.0 | 172 | 7 | 019773 |
| 13 | 29 | 100.0 | 175 | 7 | 029694 |
| 14 | 29 | 100.0 | 180 | 7 | 019607 |
| 15 | 29 | 100.0 | 180 | 7 | 019608 |
| 16 | 29 | 100.0 | 180 | 7 | 019609 |
| 17 | 29 | 100.0 | 180 | 7 | 019610 |
| 18 | 29 | 100.0 | 180 | 7 | 019611 |
| 19 | 29 | 100.0 | 180 | 7 | 019612 |
| 20 | 29 | 100.0 | 180 | 7 | 019613 |

| | | | | | | |
|----|----|-------|-----|---|--------|-------------|
| 21 | 29 | 100.0 | 181 | 7 | 046703 | homo sapien |
| 22 | 29 | 100.0 | 181 | 7 | 062917 | homo sapien |
| 23 | 29 | 100.0 | 181 | 7 | 062892 | homo sapien |
| 24 | 29 | 100.0 | 181 | 7 | 062899 | homo sapien |
| 25 | 29 | 100.0 | 181 | 7 | 062920 | homo sapien |
| 26 | 29 | 100.0 | 181 | 7 | 062922 | homo sapien |
| 27 | 29 | 100.0 | 181 | 7 | 062923 | homo sapien |
| 28 | 29 | 100.0 | 181 | 7 | 019623 | homo sapien |
| 29 | 29 | 100.0 | 181 | 7 | 019623 | homo sapien |
| 30 | 29 | 100.0 | 181 | 7 | 019747 | homo sapien |
| 31 | 29 | 100.0 | 181 | 7 | 029667 | homo sapien |
| 32 | 29 | 100.0 | 181 | 7 | 030198 | homo sapien |
| 33 | 29 | 100.0 | 181 | 7 | 029708 | homo sapien |
| 34 | 29 | 100.0 | 181 | 7 | 019631 | homo sapien |
| 35 | 29 | 100.0 | 181 | 7 | 019769 | homo sapien |
| 36 | 29 | 100.0 | 181 | 7 | 029724 | homo sapien |
| 37 | 29 | 100.0 | 181 | 7 | 029910 | homo sapien |
| 38 | 29 | 100.0 | 181 | 7 | 029679 | homo sapien |
| 39 | 29 | 100.0 | 181 | 7 | 019521 | homo sapien |
| 40 | 29 | 100.0 | 181 | 7 | 019597 | homo sapien |
| 41 | 29 | 100.0 | 181 | 7 | 029909 | homo sapien |
| 42 | 29 | 100.0 | 181 | 7 | 029701 | homo sapien |
| 43 | 29 | 100.0 | 181 | 7 | 029841 | homo sapien |
| 44 | 29 | 100.0 | 181 | 7 | 019354 | gorilla gor |
| 45 | 29 | 100.0 | 181 | 7 | 029765 | homo sapien |

ALIGNMENTS

RESULT 1
019569 PRELIMINARY; PRT; 89 AA.
AC 019569;
DI 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE MHC CLASS I ANTIGEN (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CAO K., BURDET L., ZHANG G., FERNANDEZ-VINA M.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF017320; AAB/0286.2; -.
KW MHC.
FT NON_TER 1 1
FT NON_TER 89 89
SQ SEQUENCE 89 AA; 10606 MW; 99D11089 CRC32;

Query Match 100.0%; Score 29; DB 7; Length 89;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
Db 78 RIALRY 83

RESULT 2
046697 PRELIMINARY; PRT; 90 AA.
AC 046697;
DI 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE MHC CLASS I ANTIGEN HLA-H ORTHOLOG (FRAGMENT).
GN HLA-H.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SHAMBA;
 RA GRIMSLEY C., MATHER K.A., OBER C.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF022172; AAC99794.1; -
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 90 90
 SQ SEQUENCE 90 AA; 10689 MW; 5E5F2495 CRC32;

Query Match 100.0%; Score 29; DB 7; Length 90;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 Db 79 RIALRY 84

RESULT 3
 ID 019189 PRELIMINARY; PRT; 133 AA.
 AC 019189;
 DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
 DE MHC CLASS I HISTOCOMPATIBILITY ANTIGEN-B (HLA-B-27KSH) (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEUKOCYTE;
 RA PETERSDORF E.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U18659; AAB60357.1; -
 DR MIM; 142830; -
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC I.
 FT NON_TER 1 1
 FT NON_TER 133 133
 SQ SEQUENCE 133 AA; 15491 MW; 3A3BC802 CRC32;

Query Match 100.0%; Score 29; DB 7; Length 133;
 Best Local Similarity 100.0%; Pred. No. 7.3;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 Db 31 RIALRY 36

RESULT 4
 ID 078209 PRELIMINARY; PRT; 138 AA.
 AC 078209;
 DT 01-NOV-1998 (TEMBLrel. 08, Created)
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
 DT 01-MAY-1999 (TEMBLrel. 10, Last annotation update)
 DE HUMAN LEUKOCYTE ANTIGEN PRECURSOR (FRAGMENT).
 GN HLA-A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 9800772.

RA LAFORET M., FROELICH N., PARISSADIAS A., BAUSINGER H., PFEIFFER B.,
 RA TONGIO M.M.;
 RT "An intronic mutation responsible for a low level of expression of an
 RT HLA-A*24 allele.";
 RL Tissue Antigens 50:340-346(1997).
 DR EMBL; 272423; CAA96533.1; -
 DR PFAM; PF00129; MHC_I; 1.
 KW Signal; MHC.
 FT SIGNAL 1 24
 FT NON_TER 138 138
 SQ SEQUENCE 138 AA; 15610 MW; B8417FA0 CRC32;

Query Match 100.0%; Score 29; DB 7; Length 138;
 Best Local Similarity 100.0%; Pred. No. 7.6;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 Db 103 RIALRY 108

RESULT 5
 ID 019770 PRELIMINARY; PRT; 172 AA.
 AC 019770;
 DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE MHC CLASS I HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGRONG E.,
 RA BEJCHANDRA S., JUJI T., TOKUNAGA K.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U90421; AAB50144.1; -
 DR EMBL; U90420; AAB50144.1; JOINED.
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 172 172
 SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 100.0%; Score 29; DB 7; Length 172;
 Best Local Similarity 100.0%; Pred. No. 9.7;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
 Db 72 RIALRY 77

RESULT 6
 ID 019774 PRELIMINARY; PRT; 172 AA.
 AC 019774;
 DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE MHC CLASS I HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGRONG E.,
 RA BEJCHANDRA S., BLASZYK R., GROSSE-WILDE H.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U0423; AAB50145.1; -
DR EMBL; U0422; AAB50145.1; JOINED.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.

FT NON_TER 1 1
172 172
SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 100.0%; Score 29; DB 7; Length 172;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
Db 72 RIALRY 77

RESULT 7
O19775 PRELIMINARY; PRT; 172 AA.

AC O19775;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYINGTONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGRONG E.,
RA BEJCHANDRA S.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; U0425; AAB50146.1; -
DR EMBL; U0424; AAB50146.1; JOINED.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.

FT NON_TER 1 1
172 172
SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 100.0%; Score 29; DB 7; Length 172;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
Db 72 RIALRY 77

RESULT 8
O19780 PRELIMINARY; PRT; 172 AA.

AC O19780;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYINGTONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGRONG E.,
RA BEJCHANDRA S., JUJI T., TOKUNAGA K.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U0419; AAB50143.1; -
DR EMBL; U0418; AAB50143.1; JOINED.

DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
172 172
SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 100.0%; Score 29; DB 7; Length 172;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
Db 72 RIALRY 77

RESULT 9
Q95364 PRELIMINARY; PRT; 172 AA.

AC Q95364;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE MHC HLA-B*51 PROTEIN (FRAGMENT).
GN HLA-B*51FA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]
RP SEQUENCE FROM N.A.
RA BLASZYK R.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X96473; CAA65327.1; -
DR PFAM; PF00129; MHC_I; 1.
KW MHC.

FT NON_TER 1 1
172 172
SQ SEQUENCE 172 AA; 19942 MW; 1A73E47D CRC32;

Query Match 100.0%; Score 29; DB 7; Length 172;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
Db 69 RIALRY 74

RESULT 10
O19771 PRELIMINARY; PRT; 172 AA.

AC O19771;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYINGTONG D., SIRIKONG M., LONGTA K., SRINAK D.,
RA SIRICORNT U., RUNGRONG E., BEJCHANDRA S.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U0612; AAB50151.1; -
DR EMBL; U0611; AAB50151.1; JOINED.

DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
172 172
SQ SEQUENCE 172 AA; 20026 MW; 4D9A1043 CRC32;

Best Local Similarity 100.0%; Score 29; DB 7; Length 172;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match
Best Local Similarity 100.0%; Score 29; DB 7; Length 172;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIALRY 6
Db 72 RIALRY 77

RESULT 11
ID 019772 PRELIMINARY; PRT; 172 AA.
AC 019772;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D.,
RA SIRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90614; AAB50244.1; -.
DR EMBL; U90613; AAB50244.1; JOINED.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 20026 MW; 4D9A1043 CRC32;

Query Match
Best Local Similarity 100.0%; Score 29; DB 7; Length 172;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIALRY 6
Db 72 RIALRY 77

RESULT 12
ID 019773 PRELIMINARY; PRT; 172 AA.
AC 019773;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D.,
RA SIRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90616; AAB50245.1; -.
DR EMBL; U90615; AAB50245.1; JOINED.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 20052 MW; F6214671 CRC32;

Query Match
Best Local Similarity 100.0%; Score 29; DB 7; Length 172;

Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIALRY 6
Db 72 RIALRY 77

RESULT 13
ID Q29694 PRELIMINARY; PRT; 175 AA.
AC Q29694;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE MHC CLASS I HLA-B ANTIGEN (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA PETERSDORF E.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U28759; AAB60367.1; -.
DR HSSP; P10318; IROG.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
FT VARIANT 67 67 Y -> D.
FT VARIANT 73 73 I -> T.
FT NON_TER 175 175
SQ SEQUENCE 175 AA; 20332 MW; 83A0C5C3 CRC32;

Query Match
Best Local Similarity 100.0%; Score 29; DB 7; Length 175;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIALRY 6
Db 72 RIALRY 77

RESULT 14
ID Q19607 PRELIMINARY; PRT; 180 AA.
AC Q19607;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE MHC CLASS I HLA-A (FRAGMENT).
GN HLA-A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
RA RUNGROUNG E., BEJCHANDRA S.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF030920; AAB87056.1; -.
DR EMBL; AF030919; AAB87056.1; JOINED.
DR HSSP; P01891; ITMC.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 180 180
SQ SEQUENCE 180 AA; 20811 MW; CECC3537 CRC32;

Query Match
Best Local Similarity 100.0%; Score 29; DB 7; Length 180;

Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
|
|
|
|
|
|
Db 78 RIALRY 83

RESULT 15

O19608 PRELIMINARY; PRT; 180 AA.
AC O19608;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE MHC CLASS I HLA-A (FRAGMENT).
GN HLA-A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
RA RUNGRONG E., BEJCHANDRA S.;
RL Submitted (OCT-1997) to the EMBL/Genbank/DDBJ databases.
DR EMBL; AF030922; AAB87057.1; -
DR EMBL; AF030921; AAB87057.1; JOINED.
DR HSP; P01891; 1TMC.
DR PFW; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 180
SQ SEQUENCE 180 AA; 20811 MW; CECC3537 CRC32;

Query Match 100.0%; Score 29; DB 7; Length 180;

Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIALRY 6
|
|
|
|
|
|
Db 78 RIALRY 83

Search completed: February 8, 2000, 13:17:28
Job time: 32477 sec

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-4 to: GenEmbl.* out_format : pfs

Date: Feb 8, 2000 4:37 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-Q/cgml1/uspro.spool/us08653294/runat_04022000_160701_15779/app_query.fasta.1
-DB-GenEmbl -QFMT-fastap -SUFFIX-ige -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DLEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DICALIGN=200 -THR_SCORE=pct -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM-ext -MINLEN=0 -MAXLEN=1000000 -USER-US08653294
-NCPU=6 -ICPU=3 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-4

Query length: 6

Database: GenEmbl.*

Database sequences: 821193

Database length: 1518192014

Search time (sec): 11370.480000

score_list:

| Sequence | Strd | Orig | ZScore | Escore | Len | Documentation |
|-------------------|------|-------|--------|--------|-----|----------------------------------|
| gb_pat:A63076 | + | 29.00 | 155.57 | 1.21 | 18 | A63076 Sequence 3 from Patent WO |
| gb_pat:AR049398 | + | 29.00 | 155.57 | 1.21 | 18 | AR049398 Sequence 13 from patent |
| gb_pat:I21932 | + | 29.00 | 155.57 | 1.21 | 18 | I21932 Sequence 13 from patent |
| gb_pr1:HS5108HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y08692 H.sapiens HLA-B gene, ex |
| gb_pr1:HS522HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y08693 H.sapiens HLA-B gene, ex |
| gb_pr1:HS523HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y08694 H.sapiens HLA-B gene, ex |
| gb_pr1:HS524HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09420 Human cell line THAI DCH |
| gb_pr1:HS525HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09611 Human cell line THAI DCH |
| gb_pr1:HS526HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09613 Human cell line THAI DCH |
| gb_pr1:HS527HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09615 Human cell line THAI DCH |
| gb_pr1:HS528HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09617 Human cell line THAI DCH |
| gb_pr1:HS529HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09619 Human cell line THAI DCH |
| gb_pr1:HS530HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09621 Human cell line THAI DCH |
| gb_pr1:HS531HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09623 Human cell line THAI DCH |
| gb_pr1:HS532HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09625 Human cell line THAI DCH |
| gb_pr1:HS533HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09627 Human cell line THAI DCH |
| gb_pr1:HS534HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09629 Human cell line THAI DCH |
| gb_pr1:HS535HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09631 Human cell line THAI DCH |
| gb_pr1:HS536HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09633 Human cell line THAI DCH |
| gb_pr1:HS537HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09635 Human cell line THAI DCH |
| gb_pr1:HS538HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09637 Human cell line THAI DCH |
| gb_pr1:HS539HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09639 Human cell line THAI DCH |
| gb_pr1:HS540HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09641 Human cell line THAI DCH |
| gb_pr1:HS541HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09643 Human cell line THAI DCH |
| gb_pr1:HS542HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09645 Human cell line THAI DCH |
| gb_pr1:HS543HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09647 Human cell line THAI DCH |
| gb_pr1:HS544HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09649 Human cell line THAI DCH |
| gb_pr1:HS545HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09651 Human cell line THAI DCH |
| gb_pr1:HS546HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09653 Human cell line THAI DCH |
| gb_pr1:HS547HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09655 Human cell line THAI DCH |
| gb_pr1:HS548HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09657 Human cell line THAI DCH |
| gb_pr1:HS549HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09659 Human cell line THAI DCH |
| gb_pr1:HS550HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09661 Human cell line THAI DCH |
| gb_pr1:HS551HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09663 Human cell line THAI DCH |
| gb_pr1:HS552HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09665 Human cell line THAI DCH |
| gb_pr1:HS553HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09667 Human cell line THAI DCH |
| gb_pr1:HS554HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09669 Human cell line THAI DCH |
| gb_pr1:HS555HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09671 Human cell line THAI DCH |
| gb_pr1:HS556HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09673 Human cell line THAI DCH |
| gb_pr1:HS557HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09675 Human cell line THAI DCH |
| gb_pr1:HS558HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09677 Human cell line THAI DCH |
| gb_pr1:HS559HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09679 Human cell line THAI DCH |
| gb_pr1:HS560HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09681 Human cell line THAI DCH |
| gb_pr1:HS561HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09683 Human cell line THAI DCH |
| gb_pr1:HS562HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09685 Human cell line THAI DCH |
| gb_pr1:HS563HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09687 Human cell line THAI DCH |
| gb_pr1:HS564HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09689 Human cell line THAI DCH |
| gb_pr1:HS565HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09691 Human cell line THAI DCH |
| gb_pr1:HS566HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09693 Human cell line THAI DCH |
| gb_pr1:HS567HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09695 Human cell line THAI DCH |
| gb_pr1:HS568HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09697 Human cell line THAI DCH |
| gb_pr1:HS569HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09699 Human cell line THAI DCH |
| gb_pr1:HS570HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09701 Human cell line THAI DCH |
| gb_pr1:HS571HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09703 Human cell line THAI DCH |
| gb_pr1:HS572HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09705 Human cell line THAI DCH |
| gb_pr1:HS573HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09707 Human cell line THAI DCH |
| gb_pr1:HS574HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09709 Human cell line THAI DCH |
| gb_pr1:HS575HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09711 Human cell line THAI DCH |
| gb_pr1:HS576HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09713 Human cell line THAI DCH |
| gb_pr1:HS577HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09715 Human cell line THAI DCH |
| gb_pr1:HS578HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09717 Human cell line THAI DCH |
| gb_pr1:HS579HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09719 Human cell line THAI DCH |
| gb_pr1:HS580HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09721 Human cell line THAI DCH |
| gb_pr1:HS581HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09723 Human cell line THAI DCH |
| gb_pr1:HS582HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09725 Human cell line THAI DCH |
| gb_pr1:HS583HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09727 Human cell line THAI DCH |
| gb_pr1:HS584HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09729 Human cell line THAI DCH |
| gb_pr1:HS585HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09731 Human cell line THAI DCH |
| gb_pr1:HS586HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09733 Human cell line THAI DCH |
| gb_pr1:HS587HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09735 Human cell line THAI DCH |
| gb_pr1:HS588HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09737 Human cell line THAI DCH |
| gb_pr1:HS589HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09739 Human cell line THAI DCH |
| gb_pr1:HS590HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09741 Human cell line THAI DCH |
| gb_pr1:HS591HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09743 Human cell line THAI DCH |
| gb_pr1:HS592HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09745 Human cell line THAI DCH |
| gb_pr1:HS593HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09747 Human cell line THAI DCH |
| gb_pr1:HS594HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09749 Human cell line THAI DCH |
| gb_pr1:HS595HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09751 Human cell line THAI DCH |
| gb_pr1:HS596HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09753 Human cell line THAI DCH |
| gb_pr1:HS597HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09755 Human cell line THAI DCH |
| gb_pr1:HS598HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09757 Human cell line THAI DCH |
| gb_pr1:HS599HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09759 Human cell line THAI DCH |
| gb_pr1:HS600HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09761 Human cell line THAI DCH |
| gb_pr1:HS601HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09763 Human cell line THAI DCH |
| gb_pr1:HS602HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09765 Human cell line THAI DCH |
| gb_pr1:HS603HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09767 Human cell line THAI DCH |
| gb_pr1:HS604HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09769 Human cell line THAI DCH |
| gb_pr1:HS605HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09771 Human cell line THAI DCH |
| gb_pr1:HS606HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09773 Human cell line THAI DCH |
| gb_pr1:HS607HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09775 Human cell line THAI DCH |
| gb_pr1:HS608HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09777 Human cell line THAI DCH |
| gb_pr1:HS609HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09779 Human cell line THAI DCH |
| gb_pr1:HS610HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09781 Human cell line THAI DCH |
| gb_pr1:HS611HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09783 Human cell line THAI DCH |
| gb_pr1:HS612HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09785 Human cell line THAI DCH |
| gb_pr1:HS613HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09787 Human cell line THAI DCH |
| gb_pr1:HS614HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09789 Human cell line THAI DCH |
| gb_pr1:HS615HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09791 Human cell line THAI DCH |
| gb_pr1:HS616HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09793 Human cell line THAI DCH |
| gb_pr1:HS617HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09795 Human cell line THAI DCH |
| gb_pr1:HS618HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09797 Human cell line THAI DCH |
| gb_pr1:HS619HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09799 Human cell line THAI DCH |
| gb_pr1:HS620HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09801 Human cell line THAI DCH |
| gb_pr1:HS621HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09803 Human cell line THAI DCH |
| gb_pr1:HS622HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09805 Human cell line THAI DCH |
| gb_pr1:HS623HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09807 Human cell line THAI DCH |
| gb_pr1:HS624HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09809 Human cell line THAI DCH |
| gb_pr1:HS625HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09811 Human cell line THAI DCH |
| gb_pr1:HS626HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09813 Human cell line THAI DCH |
| gb_pr1:HS627HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09815 Human cell line THAI DCH |
| gb_pr1:HS628HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09817 Human cell line THAI DCH |
| gb_pr1:HS629HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09819 Human cell line THAI DCH |
| gb_pr1:HS630HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09821 Human cell line THAI DCH |
| gb_pr1:HS631HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09823 Human cell line THAI DCH |
| gb_pr1:HS632HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09825 Human cell line THAI DCH |
| gb_pr1:HS633HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09827 Human cell line THAI DCH |
| gb_pr1:HS634HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09829 Human cell line THAI DCH |
| gb_pr1:HS635HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09831 Human cell line THAI DCH |
| gb_pr1:HS636HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09833 Human cell line THAI DCH |
| gb_pr1:HS637HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09835 Human cell line THAI DCH |
| gb_pr1:HS638HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09837 Human cell line THAI DCH |
| gb_pr1:HS639HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09839 Human cell line THAI DCH |
| gb_pr1:HS640HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09841 Human cell line THAI DCH |
| gb_pr1:HS641HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09843 Human cell line THAI DCH |
| gb_pr1:HS642HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09845 Human cell line THAI DCH |
| gb_pr1:HS643HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09847 Human cell line THAI DCH |
| gb_pr1:HS644HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09849 Human cell line THAI DCH |
| gb_pr1:HS645HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09851 Human cell line THAI DCH |
| gb_pr1:HS646HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09853 Human cell line THAI DCH |
| gb_pr1:HS647HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09855 Human cell line THAI DCH |
| gb_pr1:HS648HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09857 Human cell line THAI DCH |
| gb_pr1:HS649HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09859 Human cell line THAI DCH |
| gb_pr1:HS650HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09861 Human cell line THAI DCH |
| gb_pr1:HS651HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09863 Human cell line THAI DCH |
| gb_pr1:HS652HLA1 | + | 29.00 | 132.60 | 23.07 | 250 | Y09 |

Align seg 1/1 to: AR049398 from: 1 to: 18

1 ArgillealaleuArgTyr 6
|||||
1 CGGATCGCGCTCGCTAC 18

seq_name: gb_pat: I21932

seq_documentation_block: 18 bp DNA PAT 07-OCT-1996

LOCUS I21932
DEFINITION Sequence 13 from patent US 5525492.

ACCESSION I21932
VERSION I21932.1 GI:1602286

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS Hill, A.V.S.

TITLE Process for amplifying HLA sequences

JOURNAL Patent: US 5525492-A 13 11-JUN-1996;

FEATURES

source

1. 18

/organism="unknown"

BASE COUNT 2 a 8 c 5 g 3 t

ORIGIN

alignment_scores:

Quality: 29.00

Ratio: 4.833

Length: 6

Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x I21932

Align seg 1/1 to: I21932 from: 1 to: 18

1 ArgillealaleuArgTyr 6

|||||

1 CGGATCGCGCTCGCTAC 18

seq_name: gb_pri: HSHLABHBA

seq_documentation_block: 250 bp DNA PRI 10-OCT-1996

LOCUS HSHLABHBA

DEFINITION H.sapiens HLA-B gene, exon 2, HB(a) allele.

ACCESSION Y08692

VERSION Y08692.1 GI:1619287

KEYWORDS HLA-B gene; human leukocyte antigen.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

/tissue_type="blood"
/cell_type="white"
/lab_host="E.coli TGI"
/clone="CHBC1(a)"
/clone="CHBC2(A)"
/clone="CHBC4(a)"
14. .250
/gene="HLA-B"
/gene="HLA-B"
/note="allele HB(a)"
/number=2

gene 54 a 78 c 85 g 33 t
exon
BASE COUNT
ORIGIN

alignment_scores:

Quality: 29.00

Ratio: 4.833

Length: 6

Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x HSHLABHBA

Align seg 1/1 to: HSHLABHBA from: 1 to: 250

1 ArgillealaleuArgTyr 6

|||||

221 CGGATCGCGCTCGCTAC 238

seq_name: gb_pri: HSHLABHBB

seq_documentation_block:

LOCUS HSHLABHBB

DEFINITION H.sapiens HLA-B gene, exon 2, HB(b) allele.

ACCESSION Y08693

VERSION Y08693.1 GI:1619288

KEYWORDS HLA-B gene; human leukocyte antigen.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

BASE COUNT 51 a 78 c 87 g 34 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x HSHLABHBB

Align seg 1/1 to: HSHLABHBB from: 1 to: 250

1 ArgileAlaLeuArgTyr 6
|||||
221 CGGATCGCGCTCGCTAC 238

seq_name: gb_pr1:HSHLABHBD

seq_documentation_block:
LOCUS HSHLABHBD 250 bp DNA 10-OCT-1996 PRI
DEFINITION H.sapiens HLA-B gene, exon 2, HB(d) allele.
ACCESSION Y08694
VERSION Y08694.1 GI:1619289
KEYWORDS HLA-B gene; human leukocyte antigen.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominiidae; Homo.
AUTHORS Rojas-Munoz, A., Mendez, I. and Yunis, I.
TITLE Molecular evolution of HLA-B locus in a small population amerindian
community: The Nukak-Maku
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 250)
AUTHORS Rojas-Munoz, A.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1996) A. Rojas-Munoz, National Institute Of
Health, Immunogenetics, Avenida El Dorado, Carrera 50, Santafe De

FEATURES
source

1. .250 Location/Qualifiers
/organism="Homo sapiens"
/isolate="Norman-51"
/isolate="from amerindian community Nukak-Maku"
/db_xref="taxon:9606"
/chromosome="6"
/dev_stage="adult"
/tissue_type="blood"
/cell_type="white"
/lab_host="E.coli TGI"
/clone="CHC1(d)"
14. .250
/gene="HLA-B"
<14. .>250
/gene="HLA-B"
/note="allel HB(d)"
/number=2

gene

exon

BASE COUNT 58 a 78 c 79 g 35 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x HSHLABHBD

Align seg 1/1 to: HSHLABHBD from: 1 to: 250

1 ArgileAlaLeuArgTyr 6
|||||
221 CGGATCGCGCTCGCTAC 238

seq_name: gb_pr2:HSHLABB1

seq_documentation_block:
LOCUS HSHLABB1 250 bp DNA 22-MAR-1997 PRI
DEFINITION Human cell line THAI DCH010 MHC class I HLA-B gene (allele
HLA-B*513), exon 2.
ACCESSION U90420
VERSION U90420.1 GI:1905830
KEYWORDS
SEGMENT
SOURCE i of 2
human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 250)

AUTHORS Chandanayingyong, D., Sirikong, M., Longta, K., Srinak, D.,
Rungroung, E., Bejchandra, S., Juji, T. and Tokunaga, K.
TITLE B15 alleles (B*513)

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 250)

AUTHORS Chandanayingyong, D., Sirikong, M., Longta, K., Srinak, D.,
Rungroung, E., Bejchandra, S., Juji, T. and Tokunaga, K.
TITLE Direct Submission

JOURNAL Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand

FEATURES Location/Qualifiers

1. .250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="6p21"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH010"
1. .250
/gene="HLA-B"
/number=2
/product="MHC class I HLA-B"
55 a 83 c 80 g 32 t

BASE COUNT
ORIGIN

alignment_scores:

Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x HSHLABB1

Align seg 1/1 to: HSHLABB1 from: 1 to: 250

1 ArgileAlaLeuArgTyr 6
|||||
214 CGGATCGCGCTCGCTAC 231

seq_name: gb_pr2:HSHLABD1

seq_documentation_block:
LOCUS HSHLABD1 250 bp DNA 22-MAR-1997 PRI
DEFINITION Human cell line THAI DCH010 MHC class I HLA-B gene (allele
HLA-B*51V), exon 2.
ACCESSION U90611
VERSION U90611.1 GI:1905865
KEYWORDS
SEGMENT
SOURCE i of 2
human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

```

REFERENCE 1 (bases 1 to 250)
AUTHORS Chananayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
B*51V alleles
TITLE Unpublished
JOURNAL Chananayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
AUTHORS Siriboonrit,U., Rungroung,E. and Bejchandra,S.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES Location/Qualifiers
source 1..250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/haplotype="(b)HLA-A2,B77,Cw8N,DR12(DR52),DQ7/(c)A11.1,
B51V,Cw14,DR12(DR52),DQ1"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH010"
1..250
/genes="HLA-B"
/notes="Allele: HLA-B*51V; similar to exon 2 of B*5105"
/number=2
BASE COUNT 56 a 82 c 80 g 32 t
ORIGIN
exon
alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-4 x HSHLABD1 ..
Align seg 1/1 to: HSHLABD1 from: 1 to: 250
1 ArgilleAlaLeuArgTyr 6
|||||
214 CGGATCGCGCTCCGCTAC 231
seq_name: gb_pr2:HSHLABD1
seq_documentation_block: 250 bp DNA PRI 25-MAR-1997
LOCUS HSHLABD1
DEFINITION Human cell line THAI DCH028 MHC class I HLA-B gene (allele
HLA-B*51V), exon 2.
ACCESSION U90613
VERSION U90613.1 GI:1906033
KEYWORDS
SEGMENT 1 of 2
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 250)
AUTHORS Chananayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
B*51V alleles
TITLE Unpublished
JOURNAL Chananayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
AUTHORS Siriboonrit,U., Rungroung,E. and Bejchandra,S.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES Location/Qualifiers
source 1..250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/haplotype="(a)HLA-A11,B51V,Cw14,DR6(DR52),DQ1/(b)A2,B77,
Cw8N,DR4(DR53),DQ4"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH028"
1..250
/genes="HLA-B"
/notes="Allele: HLA-B*51V; similar to exon 2 of B*5105"
/number=2
BASE COUNT 56 a 82 c 80 g 32 t
ORIGIN
exon
alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-4 x HSHLABF1 ..
Align seg 1/1 to: HSHLABF1 from: 1 to: 250
1 ArgilleAlaLeuArgTyr 6
|||||
214 CGGATCGCGCTCCGCTAC 231
seq_name: gb_pr2:HSHLABF1
seq_documentation_block: 250 bp DNA PRI 25-MAR-1997
LOCUS HSHLABF1
DEFINITION Human cell line THAI DCH028 MHC class I HLA-B gene (allele
HLA-B*51V), exon 2.
ACCESSION U90613
VERSION U90613.1 GI:1906033
KEYWORDS
SEGMENT 1 of 2
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 250)
AUTHORS Chananayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
B*51V alleles
TITLE Unpublished
JOURNAL Chananayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
AUTHORS Siriboonrit,U., Rungroung,E. and Bejchandra,S.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES Location/Qualifiers

```

```

source 1..250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/haplotype="(a)HLA-A1,B51V,Cw14,DR6(DR52),DQ1/(b)A2,B77,
Cw8N,DR4(DR53),DQ4"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH028"
1..250
/genes="HLA-B"
/notes="Allele: HLA-B*51V; similar to exon 2 of B*5105"
/number=2
BASE COUNT 56 a 82 c 80 g 32 t
ORIGIN
exon
alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-4 x HSHLABF1 ..
Align seg 1/1 to: HSHLABF1 from: 1 to: 250
1 ArgilleAlaLeuArgTyr 6
|||||
214 CGGATCGCGCTCCGCTAC 231
seq_name: gb_pr2:HSHLABG1
seq_documentation_block: 250 bp DNA PRI 25-MAR-1997
LOCUS HSHLABG1
DEFINITION Human cell line THAI DCH011 MHC class I HLA-B gene (allele
HLA-B*51V), exon 2.
ACCESSION U90615
VERSION U90615.1 GI:1906037
KEYWORDS
SEGMENT 1 of 2
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 250)
AUTHORS Chananayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
B*51V alleles
TITLE Unpublished
JOURNAL Chananayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
AUTHORS Siriboonrit,U., Rungroung,E. and Bejchandra,S.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES Location/Qualifiers
source 1..250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/haplotype="(a)HLA-A11,B51V,Cw14,DR12(DR52),DQ7/(b)A24,
B77,Cw8N,DR12(DR52)DQ7"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH011"
1..250
/genes="HLA-B"
/notes="Allele: HLA-B*51V; similar to exon 2 of B*5105"
/number=2
BASE COUNT 56 a 82 c 80 g 32 t
ORIGIN
exon

```

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x HSHLABG1 ..
 Align seg 1/1 to: HSHLABG1 from: 1 to: 250
 1 ArgilleAlaLeuArgTyr 6
 |||||
 214 CGGATCGCGCTCGCTAC 231

seq_name: gb_pr2: HSHLABI1

seq_documentation_block:
 LOCUS HSHLABI1 250 bp DNA 22-MAR-1997
 DEFINITION Human cell line THAI DCH011 MHC class I HLA-B gene (allele
 HLA-B*1513), exon 2.
 ACCESSION U90422
 VERSION U90422.1 GI:1905834
 KEYWORDS
 SEGMENT
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
 Rungroung,E., Bejchandra,S., Blasczyk,R. and Grosse-Wilde,H.
 B15 alleles (B*1513)
 TITLE Unpublished
 JOURNAL
 REFERENCE 2 (bases 1 to 250)
 AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
 Rungroung,E., Bejchandra,S., Blasczyk,R. and Grosse-Wilde,H.
 TITLE Direct Submission
 JOURNAL Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine,
 Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
 Thailand

FEATURES

source Location/Qualifiers

1..250
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="VI"
 /map="6p21"
 /cell_type="lymphoblastoid"
 /cell_line="THAI DCH011"
 1..250
 /gene="HLA-B"
 /note="Allele: HLA-B*1513"
 /number=2
 /product="MHC class I HLA-B"
 55 a 83 c 80 g 32 t
 BASE COUNT
 ORIGIN

exon

alignment_scores:

Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x HSHLABI1 ..
 Align seg 1/1 to: HSHLABI1 from: 1 to: 250
 1 ArgilleAlaLeuArgTyr 6
 |||||
 214 CGGATCGCGCTCGCTAC 231

seq_name: gb_pr2: HSHLABJ1

seq_documentation_block:
 LOCUS HSHLABJ1 250 bp DNA 22-MAR-1997
 DEFINITION Human cell line THAI DCH028 MHC class I HLA-B gene (allele
 HLA-B*1513), exon 2.
 ACCESSION U90424
 VERSION U90424.1 GI:1905838
 KEYWORDS
 SEGMENT
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
 Rungroung,E. and Bejchandra,S.
 B15 alleles (B*1513)
 TITLE Unpublished
 JOURNAL
 REFERENCE 2 (bases 1 to 250)
 AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
 Rungroung,E. and Bejchandra,S.
 TITLE Direct Submission
 JOURNAL Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine,
 Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
 Thailand

FEATURES

source Location/Qualifiers

1..250
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="VI"
 /map="6p21"
 /cell_type="lymphoblastoid"
 /cell_line="THAI DCH028"
 1..250
 /gene="HLA-B"
 /note="Allele: HLA-B*1513"
 /number=2
 /product="MHC class I HLA-B"
 55 a 83 c 80 g 32 t
 BASE COUNT
 ORIGIN

exon

alignment_scores:

Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x HSHLABJ1 ..
 Align seg 1/1 to: HSHLABJ1 from: 1 to: 250
 1 ArgilleAlaLeuArgTyr 6
 |||||
 214 CGGATCGCGCTCGCTAC 231

seq_name: gb_pr2: HSHLABT1

seq_documentation_block:
 LOCUS HSHLABT1 250 bp DNA 22-MAR-1997
 DEFINITION Human cell line THAI DCH009 MHC class I HLA-B gene (allele
 HLA-B*1513), exon 2.
 ACCESSION U90418
 VERSION U90418.1 GI:1905826
 KEYWORDS
 SEGMENT
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,

214 CGGATCGCGCTCGCTAC 231

11

BASE COUNT 52 a 83 c 85 g 37 t 2 others
ORIGIN /pseudo

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x AF022160 ..

Align seg 1/1 to: AF022160 from: 1 to: 259

1 ArgIleAlaLeuArgTyr 6

|||||

235 CGGATCGCGCTCGCTAC 252

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-4 to: N_Geneseq_36:* out_format : pfs
Date: Feb 8, 2000 1:27 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL-frame p2n.model -DEV-rlp
-Q/cgml1/USPTO_spool/US08653294/runat_04022000_160701_15807/app_query.fasta.1
-DB-N_Geneseq_36 -QEMW-fastap -SURF1-rng -GAPOP-12.000
-GAPEXT-4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-GAPOP=6.000 -GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum62
-TRANS-human40.cdi -LIST=45 -DOALIGN=200 -THR_SCORE=pct
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM-ext -MINLEN=0
-MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT
-THREADS=1

Search information block:

Query: US-08-653-294-4
Query length: 6
Database: N_Geneseq_36:*
Database sequences: 311585
Database length: 125096042
Search time (sec): 590.520000

score_list:

| Sequence | Strid | Orig | ZScore | Escore | Len | Documentation |
|----------------------|-------|-------|--------|---------|-------|----------------------------------|
| N_Geneseq_36:Q24902 | + | 29.00 | 147.11 | 1.36 | 18 | Human leukocyte antigen probe. H |
| N_Geneseq_36:T94791 | + | 29.00 | 147.11 | 1.36 | 18 | Human leukocyte antigen class I |
| N_Geneseq_36:Q29167 | + | 29.00 | 124.48 | 24.75 | 270 | HLA-B*52 exon 2 alpha-1 domain |
| N_Geneseq_36:Q01834 | + | 29.00 | 112.84 | 110.05 | 1086 | Sequence encoding HLA-B*51 anti |
| N_Geneseq_36:Q01822 | + | 29.00 | 112.84 | 110.05 | 1086 | Sequence encoding HLA-B*52 anti |
| N_Geneseq_36:Q05693 | + | 29.00 | 112.82 | 110.38 | 1089 | HLA-B*51 gene for production of |
| N_Geneseq_36:Q05701 | + | 29.00 | 112.82 | 110.38 | 1089 | HLA-B*52 gene for production of |
| N_Geneseq_36:Q12114 | + | 29.00 | 112.82 | 110.38 | 1089 | HLA-B*53 exon. HLA-B*53 gene, |
| N_Geneseq_36:V33683 | + | 29.00 | 103.89 | 346.83 | 3168 | Candida neoformans IPC synthas |
| N_Geneseq_36:V33682 | + | 29.00 | 103.89 | 346.83 | 3168 | Candida neoformans IPC synthas |
| N_Geneseq_36:V33681 | + | 29.00 | 103.75 | 352.94 | 3220 | Candida neoformans IPC synthas |
| N_Geneseq_36:Q24977 | - | 29.00 | 99.84 | 582.73 | 5140 | DNA encoding soluble mannose x |
| N_Geneseq_36:Q74377 | - | 29.00 | 98.78 | 667.98 | 5838 | Staphylococcus aureus contig s |
| N_Geneseq_36:V51051 | + | 28.00 | 106.35 | 253.03 | 1380 | Human amine receptor cDNA. DNA |
| N_Geneseq_36:T26331 | + | 27.00 | 112.60 | 113.53 | 382 | Human gene signature HUMGS08571 |
| N_Geneseq_36:V78247 | + | 27.00 | 112.21 | 119.28 | 400 | Staphylococcus aureus contig SE |
| N_Geneseq_36:V06569 | + | 27.00 | 99.92 | 577.29 | 1741 | Rice cellulose synthase RSW1-1 |
| N_Geneseq_36:V06567 | + | 27.00 | 93.81 | 1.3e+03 | 3614 | Arabidopsis cellulose biosynth |
| N_Geneseq_36:V25775 | + | 27.00 | 89.45 | 2.2e+03 | 6093 | S. erythraea cleandomycin-synth |
| N_Geneseq_36:Q28692 | + | 26.00 | 132.77 | 8.54 | 20 | pala target oligonucleotide h. C |
| N_Geneseq_36:V23355 | + | 26.00 | 103.12 | 382.95 | 694 | Human secreted protein gene 45 |
| N_Geneseq_36:V37348 | - | 26.00 | 100.69 | 522.92 | 928 | Streptococcus pneumoniae coding |
| N_Geneseq_36:V05843 | + | 26.00 | 100.63 | 526.55 | 934 | Repeat sequence of cps gene loc |
| N_Geneseq_36:V98616 | + | 26.00 | 97.73 | 764.21 | 1322 | DNA encoding PTS system mannose |
| N_Geneseq_36:V30737 | + | 26.00 | 97.72 | 765.45 | 1324 | Streptococcus pneumoniae genom |
| N_Geneseq_36:V52429 | + | 26.00 | 97.04 | 835.08 | 1436 | Streptococcus pneumoniae genom |
| N_Geneseq_36:V14006 | + | 26.00 | 96.60 | 883.18 | 1513 | H. pylori GHP1442 gene. New |
| N_Geneseq_36:Q22665 | + | 26.00 | 96.40 | 906.98 | 1551 | Sequence of beta-tubulin gene |
| N_Geneseq_36:V767992 | + | 26.00 | 96.17 | 933.34 | 1593 | H. pylori cytoplasmic protein |
| N_Geneseq_36:V74678 | + | 26.00 | 94.93 | 1.1e+03 | 1847 | Staphylococcus aureus contig s |
| N_Geneseq_36:V79594 | - | 26.00 | 93.12 | 1.4e+03 | 2294 | Rat organic anion transporter |
| N_Geneseq_36:V23687 | + | 26.00 | 89.80 | 2.1e+03 | 3415 | L. lactis DB1341 pfl gene. Lact |
| N_Geneseq_36:V25775 | - | 26.00 | 86.47 | 3.2e+03 | 5082 | Hepatitis E virus hollow parti |
| N_Geneseq_36:V25775 | - | 26.00 | 84.96 | 3.9e+03 | 6093 | S. erythraea cleandomycin-synth |
| N_Geneseq_36:V61690 | + | 26.00 | 83.57 | 4.7e+03 | 7194 | Hepatitis E virus hollow parti |
| N_Geneseq_36:Q14412 | + | 26.00 | 83.57 | 4.7e+03 | 7195 | Forward strand of Burmese ET-N |
| N_Geneseq_36:V54729 | + | 26.00 | 83.57 | 4.7e+03 | 7195 | DNA sequence of the Burmese is |
| N_Geneseq_36:V66321 | + | 26.00 | 83.57 | 4.7e+03 | 7195 | ET-NAB (HEV) Burma strain DNA |
| N_Geneseq_36:V31256 | + | 26.00 | 76.04 | 1.2e+04 | 17710 | E. coli J96 pathogenicity isl |
| N_Geneseq_36:Q46806 | + | 26.00 | 71.66 | 2.1e+04 | 29879 | eryA region of S. erythraea c |
| N_Geneseq_36:V25726 | - | 25.00 | 123.37 | 28.53 | 36 | Primer OTG10834 for modified HPV |

N_Geneseq_36:Q33975 + 25.00 111.32 133.67 152 ! Downstream sequence of micro
N_Geneseq_36:V78694 - 25.00 105.42 285.02 308 ! Staphylococcus aureus contig
N_Geneseq_36:Q12109 - 25.00 105.26 290.98 314 ! Human pancreatic secretory t
N_Geneseq_36:V78034 - 25.00 103.24 377.21 400 ! Staphylococcus aureus contig
seq_name: N_Geneseq_36:Q24902
seq_documentation_block:
ID Q24902 standard; DNA; 18 BP.
AC Q24902;
DT 19-NOV-1992 (first entry)
DE Human leukocyte antigen probe.
KW HLA; polymerase chain reaction; inflammatory arthropathy;
KW susceptibility; arthritis; arthritis related diseases; ss.
OS Synthetic.
PN WO9207956-A.
PD 14-MAY-1992.
PF 05-NOV-1991; G01935.
PR 05-NOV-1990; GB-024005.
PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PI H11 AVS.
DR WPI; 92-183591/22.
PT PCR amplification of nucleic acids using buffer soln. and
PT chelating agent - for detecting HLA class I alleles for
PT determining susceptibility to arthritis etc.
PS Disclosure; Page 14; 52pp; English.
CC The sequence is that of a probe which hybridises to one of the
CC human leukocyte antigen (HLA) sequences in the primer extension
CC products (or strands) produced during PCR amplification of the HLA
CC class I alleles. It is specific for a sequence that encodes a Bw4
CC epitope, e.g. IALR, or the complementary nucleic acid sequence.
CC It can be used in the detection and/or identification of an HLA
CC sequence that may be indicative of a patient's susceptibility to
CC inflammatory arthropathy such as arthritis and arthritis related
CC diseases. Such diseases include reactive arthritis, rheumatoid
CC arthritis, Reiter's syndrome, uveitis, viral arthritis, psoriatic
CC arthropathy, gouty arthritis, septic arthritis, erythema nodosum,
CC Henoch-Schleien purpura and esp. ankylosing spondylitis.
CC See also Q24895-Q24901.
SQ Sequence 18 BP; 2 A; 8 C; 5 G; 3 T;

alignment_scores:

Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x Q24902 ..
Align seg 1/1 to: Q24902 from: 1 to: 18

1 ArgilleAlaLeuArgTyr 6

|||||||
1 CGGATCGCGCTCCGCTAC 18

seq_name: N_Geneseq_36:T94791

seq_documentation_block:

ID T94791 standard; DNA; 18 BP.
AC T94791;
DT 19-FEB-1998 (first entry)
DE Human leukocyte antigen class I gene URSTO probe 307-324.
KW MHC gene; major histocompatibility complex; paternity test;
KW forensic medicine; haematological malignancy; inherited disorder;
KW adoptive immunotherapy; identification; ss.
OS Synthetic.
PN WO9720197-A2.
PD 05-JUN-1997.
PR 29-NOV-1996; G02959.
PR 29-NOV-1995; GB-024381.

PA (NOLA-) NOLAN BONE MARROW TRUST ANTHONY.
 PI Arguello R. Avakian H, Madrigal A;
 DR WPI; 97-310717/28.
 PT Identifying unknown allele(s) of a polyallelic gene using panel of
 PT probes each recognising a sequence motif present in some allele(s) -
 PT useful for donor matching in tissue transplantation
 PS Claim 5; Page 19: 64pp; English.
 CC A novel method has been developed for identifying an unknown allele of a
 CC polyallelic gene. The method involves: (a) contacting the unknown allele
 CC with a panel of probes, each of which recognises a sequence motif that
 CC is present in some alleles of the polyallelic gene but not in others;
 CC (b) observing which probes recognise the unknown allele so as to obtain
 CC a fingerprint of the unknown allele; and (c) comparing the fingerprint
 CC with fingerprints of known alleles. The present sequence represents a
 CC specifically claimed probe for use in the method where the polyallelic
 CC gene is a human leukocyte antigen class I gene. The method can be used
 CC for genes such as mammalian MHC genes, specifically the HLA class I and
 CC II genes, the T cell receptor genes in mammals, TAP, LMP, ras,
 CC nonclassical HLA class I genes, human complement factor genes C4 and C2,
 CC Bf in the HLA complex, and genes located in mitochondrial DNA, bacterial
 CC chromosomes and viral DNA. The method is particularly useful for
 CC matching the alleles of the HLA genes in a prospective donor and a
 CC prospective recipient in tissue or organ transplantations. The method
 CC can also be used in paternity testing, in forensic medicine, as a
 CC follow up technique in treatment of haematological malignancies or
 CC inherited disorders, in adoptive immunotherapy, and in identification
 CC of bacteria and viruses. The method can provide for the identification
 CC of alleles of the polyallelic genes using a limited number of selected
 CC recurring motif probes.
 SQ Sequence 18 BP; 2 A; 8 C; 5 G; 3 T;

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x T94791 ..

Align seg 1/1 to: T94791 from: 1 to: 18

1 ArgilleAlaLeuArgTyr 6

1 CGGATCGCGCTCCGCTAC 18

seq_name: N_Geneseq_36:Q29167

seq_documentation_block:

ID Q29167 standard; DNA: 270 BP.

AC Q29167;

DT 09-MAR-1993 (first entry)

DE HLA-B*52 exon 2 alpha-1 domain.

KW Human leukocyte antigen; transgenic; germ cells; somatic cells;

KW expression; ss.

PN J04091731-A.

PD 25-MAR-1992.

PF 03-AUG-1990; 207329.

PR 03-AUG-1990; JP-207329.

PA (OLYU) OLYMPUS OPTICAL CO.

DR WPI; 92-342893/42.

PT Transgenic non-human mammalian HLA-B*52 gene - useful for

PT analysis of expression of gene structure, and prodn. of

PT mouse model of human disease

PS Disclosure: Fig 1; 8pp; Japanese.

CC The sequence shows the exon 2 alpha-1 domain of the human leukocyte

CC antigen-B*52 gene. The complete gene may be introduced into non-

CC human mammals, pref. rat or mouse, or their ancestors at the primary

CC developmental biological step via transplantation into the zygote or

CC embryo to generate transgenic non-human mammals incorporating the

CC HLA-B*52 gene in both their germ cells and somatic cells. Transgenic

CC non-human mammals contg. HLA-B*52 are useful for the analysis of

CC expression of the gene, its structure, and prodn. of mouse models of

CC human disease. See also Q29166-72.
 SQ Sequence 270 BP; 59 A; 88 C; 86 G; 37 T;

alignment_scores:

Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x Q29167 ..

Align seg 1/1 to: Q29167 from: 1 to: 270

1 ArgilleAlaLeuArgTyr 6

11111111111111111111

234 CGGATCGCGCTCCGCTAC 251

seq_name: N_Geneseq_36:Q01834

seq_documentation_block:

ID Q01834 standard; DNA: 1086 BP.

AC Q01834;

DT 19-MAR-1991 (first entry)

DE Sequence encoding HLA-B*51 antigen.

KW Probe: HLA class I DNA; immunogen; ss.

OS Homo sapiens.

PN EP354580-A.

PD 14-FEB-1990.

PF 10-AUG-1989.

PR 11-AUG-1988; JP-200758.

PA (OLYU) Olympus Optical Co., Ltd.

PI Kano K. Takiguchi;

DR WPI; 90-046289/07.

PT New DNA for class I human leucocyte antigens and derived probes and

PT transformed cells, useful for DNA typing, as immunogens etc.

PS Claim 1; Page 11; 23pp; English.

CC The HLA class I DNA can be used as a source of probes for use in DNA

CC typing. Transformed cells, which are useful as immunogens, can be

CC obtained by introducing these DNAs into eucaryotic cells.

SQ Sequence 1086 BP; 224 A; 334 C; 356 G; 172 T;

alignment_scores:

Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x Q01834 ..

Align seg 1/1 to: Q01834 from: 1 to: 1086

1 ArgilleAlaLeuArgTyr 6

11111111111111111111

306 CGGATCGCGCTCCGCTAC 323

seq_name: N_Geneseq_36:Q01822

seq_documentation_block:

ID Q01822 standard; DNA: 1086 BP.

AC Q01822;

DT 19-MAY-1991 (first entry)

DE Sequence encoding HLA-B*52 antigen.

KW Probe: HLA class I DNA; immunogen; ss.

OS Homo sapiens.

FN Key Location/Qualifiers

FT cds 1..1086

FT /*tag= a

PN EP-354580-A.

PD 14-FEB-1990.

PF 10-AUG-1989.

PR 11-AUG-1988; JP-200758.

PA (OLYU) Olympus Optical Co., Ltd.
PI Kano K, Takiguchi;
DR WPI: 90-046289/07.
P-PSDB; R03142.
PT New DNA for class 1 human leucocyte antigens and derived probes and
PS transformed cells, useful for DNA typing, as immunogens etc.
Claim 2; pp11-12; 23pp; English.
CC The HLA class I DNA can be used as a source of probes for use in DNA
CC typing. Transformed cells, which are useful as immunogens, can be
CC obtained by introducing these DNAs into eucaryotic cells.
SQ Sequence 1086 BP; 223 A; 335 C; 358 G; 170 T;

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x Q01822 ..
Align seg 1/1 to: Q01822 from: 1 to: 1086

1 ArgIleAlaLeuArgTyr 6
|||||
306 CGATCGCGCTCCGCTAC 323

seq_name: N_Geneseq_36:Q05693

seq_documentation_block:

ID Q05693 standard; DNA; 1089 BP.
AC Q05693;
DT 03-JAN-1991 (first entry)
DE HLA-B51 gene for production of monoclonal antibodies.
KW Allotype specific monoclonal anti-HLA antibodies; hybridomas;
KW transgenic animals; HLA-B51 gene; ss.
FH Key Location/Qualifiers
FT exon 1..73

FT exon /*tag= a
FT exon /number=1
FT exon 74..343
FT exon /*tag= b
FT exon /number=2
FT exon /note="alpha 1-domain"
FT exon 344..619
FT exon /*tag= c
FT exon /number=3
FT exon /note="alpha 2-domain"
FT exon 620..895
FT exon /*tag= d
FT exon /number=4
FT exon /note="alpha 3-domain"
FT exon 896..1012
FT exon /*tag= e
FT exon /number=5
FT exon 1013..1042
FT exon /*tag= f
FT exon /number=6
FT exon 1043..1089
FT exon /*tag= g
FT exon /number=7

EP-383183-A.
22-AUG-1990.
PD 07-FEB-1990; 102424.
PF 08-FEB-1989; JP-029313.
PR (OLYU) OLYMPUS OPTICAL KK.
PA Takiguchi M;

DR WPI: 90-255479/34.
PT Allotype specific monoclonal anti- HLA antibodies prodn. - using
PT hybridomas derived from transgenic animals carrying HLA gene and
PT immunised with HLA antigen of different allotype
PS Disclosure; Fig 1 A-G; 20pp; English.
CC The human HLA-B51 gene was injected into fertilised mouse eggs and

CC then these introduced into the uterus of a pseudo pregnant mouse.
CC The young were tested to ensure incorporation of the gene into the
CC chromosome, and one of them mated 3 times with a normal male to
CC produce 16 young, seven of which carried the HLA-B51 gene.
CC The transgenic offspring were immunised with HLA antigen.
CC The spleen lymphocytes were fused with myeloma cells. Hybridomas
CC producing antibodies were selected.
CC See also Q05701.

SQ Sequence 1089 BP; 224 A; 335 C; 357 G; 173 T;

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-4 x Q05693 ..

Align seg 1/1 to: Q05693 from: 1 to: 1089

1 ArgIleAlaLeuArgTyr 6
|||||
307 CGATCGCGCTCCGCTAC 324

seq_name: N_Geneseq_36:Q05701

seq_documentation_block:

ID Q05701 standard; DNA; 1089 BP.
AC Q05701;
DT 03-JAN-1991 (first entry)
DE HLA-Bw52 gene for production of monoclonal antibodies.
KW Allotype specific monoclonal anti-HLA antibodies; hybridomas;
KW transgenic animals; HLA-Bw52 gene; ss.
FH Key Location/Qualifiers
FT exon 1..73

FT exon /*tag= a
FT exon /number=1
FT exon 74..343
FT exon /*tag= b
FT exon /number=2
FT exon /note="alpha 1-domain"
FT exon 344..619
FT exon /*tag= c
FT exon /number=3
FT exon /note="alpha 2-domain"
FT exon 620..895
FT exon /*tag= d
FT exon /number=4
FT exon /note="alpha 3-domain"
FT exon 896..1012
FT exon /*tag= e
FT exon /number=5
FT exon 1013..1042
FT exon /*tag= f
FT exon /number=6
FT exon 1043..1089
FT exon /*tag= g
FT exon /number=7

EP-383183-A.
22-AUG-1990.
PD 07-FEB-1990; 102424.
PF 08-FEB-1989; JP-029313.
PR (OLYU) OLYMPUS OPTICAL KK.

PA Takiguchi M;
DR WPI: 90-255479/34.
PT Allotype specific monoclonal anti- HLA antibodies prodn. - using
PT hybridomas derived from transgenic animals carrying HLA gene and
PT immunised with HLA antigen of different allotype
PS Disclosure; Fig 1 A-G; 20pp; English.
CC The human HLA-Bw52 gene was introduced into mouse L cells and
CC then these cells used to immunise one of the transgenic mice
CC (See Q05693).

CC The spleen lymphocytes were fused with myeloma cells (P3x63-Ag8.653).
 CC Hybridomas producing antibodies were selected.
 SQ Sequence 1089 BP; 223 A; 336 C; 359 G; 171 T;

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x Q05701 ..

Align seg 1/1 to: Q05701 from: 1 to: 1089

1 ArgileAlaLeuArgTyr 6
 |||||
 307 CGGATCGGCTCCGCTAC 324

seq_name: N_Geneseq_36:Q12114

seq_documentation_block:

ID Q12114 standard; DNA; 1089 BP.

AC Q12114;

DT 29-AUG-1991 (first entry)

DE HLA-Bw53 exon.

KW Human leukocyte antigen; probe; major histocompatibility complex;

KW MHC; class I; ss.

OS Homo sapiens.

FH Key

FT cds

Location/Qualifiers

1..1089

/*tag- a

J03112487-A.

PD 14-MAY-1991.

PF 22-SEP-1989; 247697.

PR 22-SEP-1989; JP-247697.

PA (OLYU) OLYMPUS OPTICAL KK.

DR WPI; 91-182991/25.

P-PSDB; R12463.

PT HLA-Bw53 gene, DNA probe and transformant cells - used for

PT immunisation, identifying specificity of antiserum etc.

PS Claim 1; Page 1; lpp; Japanese.

CC Probes comprising part of the sequence can be used to identify

CC Class I genes. The DNA can be expressed for immunisation of

CC animals and prodn. of monoclonal antibodies specific for the

CC HLA-Bw53 antigen. See also J03112485 and J03112486.

SQ Sequence 1089 BP; 222 A; 337 C; 356 G; 174 T;

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x Q12114 ..

Align seg 1/1 to: Q12114 from: 1 to: 1089

1 ArgileAlaLeuArgTyr 6
 |||||
 307 CGGATCGGCTCCGCTAC 324

seq_name: N_Geneseq_36:V33683

seq_documentation_block:

ID V33683 standard; mRNA; 3168 BP.

AC V33683;

DT 19-JAN-1999 (first entry)

DE Candida neoformans IPC synthase mRNA.

KW Candida; IPC synthase; fungus; inositolphosphoryl ceramide synthase;

KW anti-fungal therapy; sphingolipid biosynthesis; phosphatidylinositol; ss.

OS Candida neoformans.

PN EP-872485-A2.

PD 21-OCT-1998.

PF 14-APR-1998; 302866.

PR 17-OCT-1997; US-062971.

PR 15-APR-1997; US-043591.

PR 22-APR-1997; US-044095.

PR 13-MAY-1997; US-046348.

PR 21-JUL-1997; US-053320.

PA (ELIL) LILLY & CO ELI.

PI Heider SA, Radding JA;

DR WPI; 98-533879/46.

PT New inositolphosphoryl ceramide synthase genes from fungi - useful

PT for identifying compounds for anti-fungal therapy

PS Claim 5; Page 49-50; 53pp; English.

CC The present sequence represents the mRNA of a pure inositolphosphoryl

CC ceramide (IPC) synthase protein from a fungal cell, Candida neoformans.

CC The present invention also describes a method for identifying inhibitory

CC compounds of fungal IPC synthase protein activity. IPC synthase proteins

CC are useful for identifying inhibitors of fungal sphingolipid

CC biosynthesis, as the IPC synthase catalyses a step in the synthesis of

CC inositolphosphoryl ceramide from ceramide and phosphatidylinositol.

CC Fragments of IPC synthase proteins are also useful as probes or primers

CC for identification and isolation of homologous genes.

SQ Sequence 3168 BP; 692 A; 899 C; 812 G; 765 U;

alignment_scores:

Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x V33683

Align seg 1/1 to: V33683 from: 1 to: 3168

1 ArgileAlaLeuArgTyr 6

|||||

2608 AGAAUAGCUCUAGAU 2625

seq_name: N_Geneseq_36:V33682

seq_documentation_block:

ID V33682 standard; cDNA; 3168 BP.

AC V33682;

DT 19-JAN-1999 (first entry)

DE Candida neoformans IPC synthase encoding cDNA.

KW Candida; IPC synthase; fungus; inositolphosphoryl ceramide synthase;

KW anti-fungal therapy; sphingolipid biosynthesis; phosphatidylinositol; ss.

OS Candida neoformans.

FH Key

Location/Qualifiers

1224..2621

/*tag- a

PN EP-872485-A2.

PD 21-OCT-1998.

PF 14-APR-1998; 302866.

PR 17-OCT-1997; US-062971.

PR 15-APR-1997; US-043591.

PR 22-APR-1997; US-044095.

PR 13-MAY-1997; US-046348.

PR 21-JUL-1997; US-053320.

PA (ELIL) LILLY & CO ELI.

PI Heider SA, Radding JA;

DR WPI; 98-533879/46.

DR P-PSDB; W70520.

PT New inositolphosphoryl ceramide synthase genes from fungi - useful

PT for identifying compounds for anti-fungal therapy

PS Claim 5; Page 46-49; 53pp; English.

CC The present sequence encodes a pure inositolphosphoryl ceramide (IPC)

CC synthase protein from a fungal cell, Candida neoformans. The present

CC invention also describes a method for identifying inhibitory compounds

CC of fungal IPC synthase protein activity. IPC synthase proteins are

CC useful for identifying inhibitors of fungal sphingolipid biosynthesis,

CC as the IPC synthase catalyses a step in the synthesis of
 CC inositolphosphoryl ceramide from ceramide and phosphatidylinositol.
 CC Fragments of IPC synthase proteins are also useful as probes or primers
 CC for identification and isolation of homologous genes.
 SQ Sequence 3168 BP; 592 A; 899 C; 812 G; 765 T;

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x V33682 ..

Align seg 1/1 to: V33682 from: 1 to: 3168

1 ArgIleAlaLeuArgTyr 6

|||||

2608 AGAATAGCTCTTAGATAT 2625

seq_name: N_Geneseq_36:V33681

seq_documentation_block:

ID V33681 standard; DNA; 3220 BP.

AC V33681;

DT 19-JAN-1999 (first entry)

DE Candida neoformans IPC synthase DNA.

KW Candida; IPC synthase; fungus; inositolphosphoryl ceramide synthase;

KW anti-fungal therapy; sphingolipid biosynthesis; phosphatidylinositol; ss.

OS Candida neoformans.

FH Key Location/Qualifiers

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

FT intron 1888..1939

seq_name: N_Geneseq_36:Q24977
 seq_documentation_block:
 ID Q24977 standard; DNA; 5140 BP.
 AC Q24977;
 DT 18-NOV-1992 (first entry)
 DE DNA encoding soluble mannose receptor peptide.
 KW Soluble mannose receptor peptide; ss DNA; receptor binding;
 KW cancer cells; targeting; probe; drug delivery; cell marker;
 KW fusion molecule; Azi; ricin; pertussis; cholera toxin; liposome;
 KW therapeutic agent; diagnostic agent; opportunistic infections;
 KW immunocompromised patients; HIV.
 OS Homo sapiens.

FH Key Location/Qualifiers

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x Q24977/rev ..

Align seg 1/1 to reverse of: Q24977 from: 1 to: 5140

1 ArgIleAlaLeuArgTyr 6

|||||

4446 AGAATCGCATTGAGTAC 4429

seq_name: N_Geneseq_36:V74377

seq_documentation_block:

ID V74377 standard; DNA; 5838 BP.

AC V74377;

DT 16-MAR-1999 (first entry)

DE Staphylococcus aureus contig SEQ ID #66.

KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.
 OS Staphylococcus aureus.
 FH Key Location/Qualifiers
 FT misc_feature 421..480
 FT /*tag= a
 FT /note= "these bases represent a line of missing text in
 the sequence listing in the specification. They
 are included to maintain the nucleotide numbering
 given in the specification for this DNA sequence"
 FT misc_feature 2221..2280
 FT /*tag= b
 FT /note= "these bases represent a line of missing text in
 the sequence listing in the specification. They
 are included to maintain the nucleotide numbering
 given in the specification for this DNA sequence"
 FT misc_feature 4021..4080
 FT /*tag= c
 FT /note= "these bases represent a line of missing text in
 the sequence listing in the specification. They
 are included to maintain the nucleotide numbering
 given in the specification for this DNA sequence"
 FT misc_feature 5821..5838
 FT /*tag= d
 FT /note= "these bases represent a line of missing text in
 the sequence listing in the specification. They
 are included to maintain the nucleotide numbering
 given in the specification for this DNA sequence"
 FT EP-786519-A2.
 PN 30-JUL-1997.
 PD 07-JAN-1997; 100117.
 PR 03-JAN-1996; US-009861.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
 PI Rosen CA;
 DR WPI: 97-374922/35.
 PT Polynucleotide(s) and proteins derived from Staphylococcus aureus
 scored on computer readable medium and used in the production of
 anti-S.aureus vaccines
 PS Claim 1; Page 488-491; 3271pp; English.
 CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
 of the invention. The DNA sequences are recorded on a computer readable
 medium, preferably selected from a floppy or hard disk, random access
 memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 the S.aureus DNA sequences allows putative functions to be assigned so
 that protein-encoding or regulatory regions of commercial, therapeutic or
 industrial importance can be obtained. Specifically, sequences which are
 likely to encode antigens have been identified and these polypeptides can
 be used in a vaccine composition against S.aureus infection. The
 polypeptides can also be used in a kit for the immunodetection of
 S.aureus in a sample. S.aureus is implicated in numerous human diseases,
 including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 skin and surgical wound infections, scalded skin syndrome, toxic shock
 syndrome, etc. Organisms transformed with the DNA sequences can be used
 for recombinant production of the polypeptides. The new DNA sequences
 (and their fragments) are useful as primers or probes for isolating
 homologues of any of the S.aureus DNA sequences contained on the
 computer readable medium.
 SQ Sequence 5838 BP; 1887 A; 979 C; 819 G; 1946 T;

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-08-653-294-4 x V74377/rev ...

Align seg 1/1 to reverse of: V74377 from: 1 to: 5838

1 ArgTleAlaLeuArgTyr 6
 |||||
 1333 CGCATTGCTTGGCATAT 1316

seq_name: N_Geneseq_36:T51051

seq_documentation_block:

ID T51051 standard; cDNA; 1380 BP.
 AC T51051;
 DT 05-APR-1997 (first entry)
 DE Human amine receptor cDNA.
 KW Amine receptor; 7-transmembrane receptor; neurotransmitter;
 KW signal transduction; gene therapy; diagnosis; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 252..1265
 FT /*tag= a
 FT primer_bind complement (252..269)
 FT /*tag= b
 FT /note= "5' primer binding site"
 FT primer_bind 1245..1265
 FT /*tag= c
 FT /note= "3' primer binding site"
 PN W09639440-A1.
 PD 12-DEC-1996.
 PF 06-JUN-1995; U07221.
 PR 06-JUN-1995; WO-U07221.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Li Y, Ruben SM;
 DR WPI: 97-043075/04.
 DR P-PSDB; W09110.
 PT DNA encoding human amine receptor - used to identify agonists and
 antagonists, or amine neurotransmitters
 PS Claim 1; Fig 1A-1C; 68pp; English.
 CC A cDNA clone (T51051) codes for a human 7-transmembrane receptor
 (W09110) putatively identified as an amine receptor on the basis
 of sequence homology to the rat amine receptor. It was discovered
 in a human genomic library and can be found in human monocytes.
 CC Isolation of the cDNA allows recombinant prodn. of human amine
 receptor in transfected host (e.g. E. coli, COS, insect) host
 cells. Polynucleotides can be used as probes to detect mutations
 in the amine receptor gene, and in the gene therapy of conditions
 CC associated with under-expression of the receptor.
 SQ Sequence 1380 BP; 288 A; 371 C; 324 G; 397 T;

alignment_scores:
 Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
 US-08-653-294-4 x T51051 ...

Align seg 1/1 to: T51051 from: 1 to: 1380

1 ArgTleAlaLeuArgTyr 6

|||||
 696 AGGTGGCTCTCAGGTAC 713

seq_name: N_Geneseq_36:T26331

seq_documentation_block:

ID T26331 standard; cDNA to mRNA; 382 BP.
 AC T26331;
 DT 16-OCT-1996 (first entry)
 DE Human gene signature HUMGS08571.
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 KW cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN W09514772-A1.
 PD 01-JUN-1995.

PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATSU) MATSUBARA K.
PA (OKUBO) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 2058-2059; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 382 BP; 118 A; 72 C; 85 G; 102 T;

alignment_scores:
Quality: 27.00 Length: 6
Ratio: 4.500 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
US-08-653-294-4 x T26331 ..

Align seg 1/1 to: T26331 from: 1 to: 382

1 ArgilleAlaLeuArgTyr 6
||||:|||||
54 CGCCTGGCCCTCAGATAT 71

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-4 to: EST:* out_format : pfs

Date: Feb 8, 2000 4:02 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters: -DEV-xlp

-Q/cgml_1/USPTO.spool/US08653294/runat_04022000_160700_15770/app_query.fasta.1
-DB-EST -QFMT-fastap -SUFFIX-rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -CGAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-XGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELEX=7.000 -START=1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST=45 -DOALIGN=200 -THRESHOLD=15 -MODE=LOCAL
-OUTFMT=pfs -NORM-ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294
-NCPU=6 -ICPU=3 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-4

Query length: 6

Database: EST*

Database sequences: 4538634

Database length: 1887831982

Search time (sec): 8553.360000

score_list:

| Sequence | Strd | Orig | ZScore | Escore | Len | Documentation |
|-------------------|------|-------|--------|--------|-----|---------------------------------|
| gb_est31:H14029 | - | 29.00 | 137.29 | 70.04 | 133 | H14029 EST00056 Chromosome 19p1 |
| gb_est8:C03945 | + | 29.00 | 132.48 | 129.59 | 232 | C03945 C03945 Human heart cDNA |
| gb_gss1:CN5000M8 | + | 29.00 | 131.87 | 140.23 | 249 | AL080978 Arabidopsis thaliana g |
| gb_est10:AA151891 | + | 29.00 | 131.67 | 144.00 | 255 | AA151891 zoolf06.r1 Stratagene |
| gb_gss8:B95928 | + | 29.00 | 131.63 | 144.62 | 256 | B95928 T1812TF TANU Arabidopsi |
| gb_est11:AA263158 | + | 29.00 | 130.77 | 161.60 | 283 | AA263158 PMY0534 KGI-a Lambda Z |
| gb_est4:AV298815 | + | 29.00 | 130.77 | 161.60 | 283 | AV298815 AV298815 RIKEN full-l |
| gb_est11:T24384 | + | 29.00 | 128.91 | 205.12 | 351 | T24384 cirs159 lambdaZAPST Rici |
| gb_est5:ID68742 | + | 29.00 | 128.69 | 210.95 | 360 | D8742 CELK0552F Yuj1 Kohara U |
| gb_est27:AI440253 | + | 29.00 | 128.38 | 219.40 | 373 | AI440253 t301c04.x1 NCI_CGAP_G |
| gb_est16:DB2221 | + | 29.00 | 128.34 | 220.70 | 375 | D8221 HUMHC4626 Human pancrea |
| gb_est18:AA710058 | + | 29.00 | 128.00 | 230.49 | 390 | AA710058 vt39g09.r1 Barstead m |
| gb_est37:AI945988 | + | 29.00 | 127.71 | 239.02 | 403 | AI945988 bs19b12.r1 Drosophila |
| gb_est20:AA817413 | + | 29.00 | 127.45 | 246.91 | 415 | AA817413 LD22688.5prime LD Dros |
| gb_est9:AA067065 | + | 29.00 | 127.16 | 256.81 | 430 | AA067065 mm31d10.r1 Stratagene |
| gb_est24:AI212811 | + | 29.00 | 127.13 | 257.47 | 431 | AI212811 y3el2ai.r1 Aspergillus |
| gb_est10:AA178090 | + | 29.00 | 127.11 | 258.13 | 432 | AA178090 mt06g07.r1 Soares mous |
| gb_gss11:AO298218 | + | 29.00 | 127.04 | 260.78 | 436 | AO298218 HS_3074_B2_B06.MF CIT |
| gb_est8:AA043388 | + | 29.00 | 126.80 | 268.74 | 448 | AA043388 zk53g12.r1 Soares preg |
| gb_gss6:AO843237 | + | 29.00 | 126.54 | 278.05 | 462 | AO843237 nbxb006f05r CUGI Rice |
| gb_est44:AV395907 | + | 29.00 | 126.28 | 287.40 | 476 | AV395907 AV395907 Chlamydomonas |
| gb_est20:AA817414 | + | 29.00 | 126.12 | 293.42 | 485 | AA817414 LD22688.5prime LD Dros |
| gb_gss1:CN5000RU | + | 29.00 | 126.03 | 296.77 | 490 | AL08192 Arabidopsis thaliana g |
| gb_est3:HO5321 | + | 29.00 | 125.99 | 298.11 | 492 | HO5321 y180h03.s1 Soares infant |
| gb_gss14:AO569105 | + | 29.00 | 125.97 | 298.78 | 493 | AO569105 HS_5189_B2_A01.T7A RPO |
| gb_gss15:AO620992 | + | 29.00 | 125.63 | 312.23 | 513 | AO620992 HS_5334_A1_F05.T7A RPO |
| gb_gss6:AO890557 | + | 29.00 | 125.60 | 313.58 | 515 | AO890557 HS_2204_A1_G03.MR CIT |
| gb_gss11:AO265576 | + | 29.00 | 125.37 | 316.28 | 519 | AO265576 RBC111-69P21.TK RPCI-1 |
| gb_gss5:AO808427 | + | 29.00 | 125.37 | 323.03 | 529 | AO808427 HS_4553_B2_G10.T7A CIT |
| gb_gss13:AO467244 | + | 29.00 | 125.17 | 331.15 | 541 | AO467244 HS_5210_A2_B02.T7A RPO |
| gb_gss10:AO186377 | + | 29.00 | 125.01 | 337.94 | 551 | AO186377 HS_3085_A1_G10.MF CIT |
| gb_gss11:AO295519 | + | 29.00 | 124.95 | 340.66 | 555 | AO295519 HS_3083_A1_H12.MF CIT |
| gb_gss4:AO899032 | + | 29.00 | 124.91 | 342.70 | 558 | AO899032 HS_5367_A2_F04.T7A RPO |
| gb_est3:HO5555 | + | 29.00 | 124.26 | 372.05 | 601 | HO5555 y175c04.r1 Soares infant |
| gb_gss15:AO658515 | + | 29.00 | 124.24 | 373.42 | 603 | AO658515 Sheared DNA-9J12.TF SH |
| gb_est18:AA695239 | + | 29.00 | 124.22 | 374.11 | 604 | AA695239 GM02463.5prime GM Dros |
| gb_est26:AT359260 | + | 29.00 | 124.02 | 383.72 | 618 | AT359260 qv27b07.x1 NCI_CGAP_B |
| gb_est21:AA950260 | + | 29.00 | 123.97 | 386.47 | 622 | AA950260 LD28753.5prime LD Dros |
| gb_gss11:AO336682 | + | 29.00 | 123.82 | 394.05 | 633 | AO336682 HS_5003_B2_H11.SP6E RF |
| gb_gss15:AO658631 | + | 29.00 | 123.68 | 400.95 | 643 | AO658631 Sheared DNA-24G22.TR S |
| gb_est21:AA942135 | + | 29.00 | 123.11 | 431.43 | 687 | AA942135 LD26186.5prime LD Dros |
| gb_est31:AI696864 | + | 29.00 | 122.38 | 474.04 | 748 | AI696864 wc74h11.x1 NCI_CGAP_P |

gb_est21:AA941096 - 29.00 122.23 483.17 761 I AA941096 LD25055.5prime LD D
gb_gss11:AO330948 - 29.00 122.05 494.43 777 I AO330948 nbxb0048020r CUGI R
gb_est35:AI828628 + 29.00 122.03 495.84 779 I AI828628 wc10f08.x1 NCI_CGAP

seq_name: gb_est3:H14029

seq_documentation_block: 133 bp DNA EST 03-JUL-1995
LOCUS H14029
DEFINITION EST00056 Chromosome 19p12-pl3.1 exon Homo sapiens genomic clone
ACCESSION H14029
VERSION H14029.1 GI:888038
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Chordata; Metazoa; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 133)

AUTHORS Li,Q.Y.

TITLE Chromosome 19p12-pl3.1 exons

JOURNAL Unpublished (1995)

COMMENT On Sep 21, 1992 this sequence version replaced gi:279277.

Contact: Li QY

Human Molecular Genetics

Queen's Medical Centre

Nottingham, NG7 2UH, UK

Tel: 1159249924

Fax: 1159709906

Email: pdzqyl@pdn1.gene.nottingham.ac.uk

Insert Length: 972 Std Error: 0.00

Seq primer: SD2 : 5' ATC TCA GTG GTA TTT GTC AGC 3'

High quality sequence stop: 331.

FEATURES

source

Location/Qualifiers

1..133

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="19p12-pl3.1"

/clones="CS-2"

/clone_id="Chromosome 19p12-pl3.1 exon"

/lib_host="E. coli DH5a"

/note="vector: PAMPI0; Exons were isolated from human

Livermore National Laboratory using a modification of the

method of exon amplification (Proc. Natl. Acad. Sci. USA

88: 4005-4009, 1991). Amplified exons were cloned into

PAMPI0 by uracil cloning (GIBCOL BRL)."

BASE COUNT 29 a 33 c 36 g 35 t

ORIGIN

alignment_scores:

Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x H14029/rev ..

Align seg 1/1 to reverse of: H14029 from: 1 to: 133

1 ArgilleAlaLeuArgTyr 6

|||||

127 CGATCGCCCTGAGATAT 110

seq_name: gb_est8:C03945

seq_documentation_block: 232 bp mRNA EST 30-JUL-1996

LOCUS C03945

DEFINITION C03945 Human heart cDNA (Ynakamura) Homo sapiens cDNA clone

ACCESSION 3NHC2454, mRNA sequence.

VERSION C03945

KEYWORDS C03945.1 GI:1467196

```

SOURCE
ORGANISM      human.
               Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Tanaka,T., Ogiwara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and
               Nakamura,Y.
TITLE        Construction of a normalized directionally cloned cDNA library from
               adult heart and analysis of 3040 clones by partial sequencing
JOURNAL      Genomics 35 (1), 231-235 (1996)
MEDLINE      96299762
COMMENT      On Oct 24, 1995 this sequence version replaced gi:1040105.
               Contact: Yusuke Nakamura
               Institute of Medical Science
               University of Tokyo
               4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
               Tel: 81-3-5449-5372
               Fax: 81-3-5449-5433
               Email: yusuke@ims.u-tokyo.ac.jp.
FEATURES
source
1..232
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone_lib="3NHC2454"
   /dev_stage="adult"
   /note="Organ: heart; normalized directionally cloned cDNA
               from adult heart"
BASE COUNT   55 a 77 c 68 g 32 t
ORIGIN
alignment_scores
  Quality: 29.00      Length: 6
  Ratio: 4.833       Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-4 x C03945
Align seg 1/1 to: C03945 from: 1 to: 232
seq_documentation_block:
LOCUS      C03945             249 bp      DNA             28-JUN-1999
DEFINITION Arabidopsis thaliana genome survey sequence SP6 end of BAC F2B11 of
               IGF library from strain Columbia of Arabidopsis thaliana, genomic
               survey sequence.
ACCESSION  AL080978
VERSION    AL080978.1 GI:5282118
KEYWORDS   GSS.
SOURCE     Arabidopsis thaliana
            thale cress.
            Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
            eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
            Arabidopsids.
REFERENCE  1 (bases 1 to 249)
AUTHORS   Salanoubat,M., Artiguenave,F., Brottier,P., Wincker,P.,
            Samson,D., Saurin,W., Weissenbach,J. and Quetier,F.
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 249)
AUTHORS   Genoscope.
TITLE     Direct Submission
JOURNAL   Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
            Location/Qualifiers
FEATURES

```

```

1..249
   /organism="Arabidopsis thaliana"
   /strain="Columbia"
   /db_xref="taxon:3702"
   /clone_lib="IGF"
   /clone="F2B11"
   /note="end : SP6"
BASE COUNT   63 a 55 c 47 g 84 t
ORIGIN
alignment_scores
  Quality: 29.00      Length: 6
  Ratio: 4.833       Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-4 x CNS00MR8
Align seg 1/1 to: CNS00MR8 from: 1 to: 249
seq_name: gb_est10:AA151891
seq_documentation_block:
LOCUS      AA151891             255 bp      mRNA             10-DEC-1996
DEFINITION zoolf06.r1 Stratagene colon (#937204) Homo sapiens cDNA clone
               IMAGE:566435 5' similar to gb:M15497_cds1 HLA CLASS I
               HISTOCOMPATIBILITY ANTIGEN, A-24(A-9) A*2401 (HUMAN);, mRNA
               sequence.
ACCESSION  AA151891
VERSION    AA151891.1 GI:1720754
KEYWORDS   EST.
SOURCE     human.
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 255)
AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chisoso,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfig,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
            Trevasakis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
            Generation and analysis of 280,000 human expressed sequence tags
            Genome Res. 6 (9), 807-828 (1996)
            97044478
            On May 8, 1995 this sequence version replaced gi:800234.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 1.
            Location/Qualifiers
            1..255
            /organism="Homo sapiens"
            /db_xref="GDB:4590888"
            /db_xref="taxon:9606"
            /clone="IMAGE:566435"
            /clone_lib="Stratagene colon (#937204)"
            /lab_host="SOLR cells (kanamycin resistant)"
            /note="Organ: colon; Vector: pluescript SK; Site:1;
            EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:

```

Oligo dt. T-84 colonic epithelial cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGG 3' -3' adaptor sequence: 5' CTCGATTTTTTTTTTTT 3'

BASE COUNT 57 a 70 c 75 g 44 t 9 others
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x AA151891 ..

Align seg 1/1 to: AA151891 from: 1 to: 255

1 ArgilleAlaLeuArgTyr 6

|||||

89 CGGATCGCCTCGCTAC 106

seq_name: gb_gss8:B95928

seq_documentation_block:

LOCUS B95928 256 bp DNA GSS 31-MAR-1998
DEFINITION T1812TF TAMU Arabidopsis thaliana genomic clone T1812, genomic survey sequence.

ACCESSION B95928

VERSION B95928.1 GI:2997866

KEYWORDS GSS.

SOURCE

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; eudicotyledons; Magnoliophyta; Magnoliopsida; core eudicotyledons; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 256)

AUTHORS Rounsley S.D., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Adams, M.D. and Venter, J.C.

TITLE A BAC End Sequence Database for Identifying Minimal Overlaps in Arabidopsis Genomic Sequencing. Update 3

JOURNAL

COMMENT

Other GSSs: T1812TR

Contact: Steve Rounsley

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: rounsley@tigr.org

Seq primer: M13-21

Class: BAC ends

High quality sequence stop: 256.

FEATURES

source

1..256

/organism="Arabidopsis thaliana"

/strain="Columbia"

/db_xref="taxon:3702"

/clone="T1812"

/clone.lib="TAMU"

/sex="hermaphrodite"

/note="vector: BelosACII; Site_1: HindIII; Site_2: HindIII; Produced by Rod Wing"

BASE COUNT

ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x B95928/rev ..

Align seg 1/1 to reverse of: B95928 from: 1 to: 256

1 ArgilleAlaLeuArgTyr 6

|||||

182 AGGATCGCCTCGCTAT 165

seq_name: gb_est11:AA263158

seq_documentation_block:

LOCUS AA263158 283 bp mRNA EST 02-JUL-1998
DEFINITION PMY0534 KGI-a Lambda Zap Express cDNA library Homo sapiens CDNA 5', mRNA sequence.

ACCESSION AA263158

VERSION AA263158.1 GI:1898964

KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 283)

Claudio, J.O., Liew, C.C., Dempsey, A.A., Cukerman, E., Stewart, A.K., Na, E., Atkins, H.I., Iscove, N.N. and Hawley, R.G.

TITLE Identification of sequence-tagged transcripts differentially expressed within the human hematopoietic hierarchy

JOURNAL Genomics 50 (1), 44-52 (1998)

MEDLINE 98292493

COMMENT

On May 5, 1995 this sequence version replaced gi:797810.

Contact: Hawley RG

Oncology Research Laboratories

The Toronto Hospital

CRCS-424, 67 College St., Toronto, Ontario M5G 2M1, Canada

Tel: 416 3403834

Fax: 416 3403453

Email: r.hawley@utoronto.ca

Similar to M58636 MHC class I HLA-B* gene. Clone was randomly

picked from KGIa primary library.

Seq primer: 5' GAATTAACCTCCTCAAGGG 3'

High quality sequence stop: 283.

FEATURES

source

1..283

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone.lib="KGI-a Lambda Zap Express cDNA library"

/cell_type="promyeloblast"

/cell_line="KGI-a"

/note="vector: Lambda Zap Express (Stratagene); Site_1: EcoRI; Site_2: XhoI; Unidirectional cloning sites: EcoRI-XhoI. mRNA was purified from KGI-a cell line. cDNA

was synthesized using an XhoI-OligodT linker primer. EcoRI

adaptors were ligated, followed by digestion with XhoI for

directional cloning into predigested Lambda Zap Express"

BASE COUNT

ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x AA263158 ..

Align seg 1/1 to: AA263158 from: 1 to: 283

1 ArgilleAlaLeuArgTyr 6

|||||

132 CGGATCGCCTCGCTAC 149

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x T24384/rev ..

Align seg 1/1 to reverse of: T24384 from: 1 to: 351

1 ArgileAlaLeuArgTyr 6

|||||

291 AGGATTCCTACGATAC 274

seq_name: gb_est5:D68742

seq_documentation_block:

LOCUS D68742 360 bp mRNA EST 07-DEC-1995
DEFINITION CELK055B2F Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA
clone_YK55B2 5', mRNA sequence.

ACCESSION D68742

VERSION D68742.1 GI:1104384

KEYWORDS

EST.

SOURCE

Caenorhabditis elegans.

ORGANISM

Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 360)
AUTHORS Kohara.Y., Mitsuki.H., Nishigaki.A., Motohashi.T., Sugimoto.A. and
Tabara.H.

TITLE Toward an expression map of the C.elegans genome

JOURNAL Unpublished (1994)

COMMENT On Apr 14, 1993 this sequence version replaced gi:785653.

Contact: Yuji Kohara

Gene Library Lab

National Institute of Genetics

Yata 1111, Mishima, Shizuoka 411, Japan

Tel: 81-559-81-6854

Fax: 81-559-81-6855

Email: ykohara@lab.nig.ac.jp

Insert Length: 910 Std Error: 0.00

High quality sequence stop: 326.

FEATURES

source

1..360

/organism="Caenorhabditis elegans"

/strain="CB1489 him-8(e1489)"

/db_xref="taxon:6239"

/clone="YK55B2"

/clone_lib="Yuji Kohara unpublished cDNA"

/sex="hermaphrodite, male"

/tissue_type="whole animal"

/dev_stage="varied"

BASE COUNT 98 a 102 c 83 g 76 t 1 others

ORIGIN

alignment_scores:

Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x D68742 ..

Align seg 1/1 to: D68742 from: 1 to: 360

1 ArgileAlaLeuArgTyr 6

|||||

321 CGTATTCCTACGATAC 338

seq_name: gb_est27:A1440253

seq_documentation_block:

LOCUS A1440253 373 bp mRNA EST 13-APR-1999
DEFINITION tJ01c04.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2140230 3'
similar to SW:NU4M_PANTR P03906 NADH-UBIQUINONE OXIDOREDUCTASE

CHAIN 4 ;, mRNA sequence.

ACCESSION A1440253

VERSION A1440253.1 GI:4281347

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 373)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

On Jun 5, 1998 this sequence version replaced gi:3189628.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert_Strausberg@nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/dbp/image/image.html

Insert Length: 1410 Std Error: 0.00

Seq Primer: -40UP from Gibco

High quality sequence stop: 97.

Location/Qualifiers

1..373

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2140230"

/clone_lib="NCI_CGAP_Gas4"

/tissue_type="poorly differentiated adenocarcinoma with

signet ring cell features"

/lab_host="DH10B"

/note="Organ: stomach; Vector: pCMV-SPORT6; Site:1: SalI;

Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.69 kb. Life Technologies catalog #:

11549-011"

BASE COUNT 82 a 92 c 115 g 84 t

ORIGIN

alignment_scores:

Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-4 x A1440253/rev ..

Align seg 1/1 to reverse of: A1440253 from: 1 to: 373

1 ArgileAlaLeuArgTyr 6

|||||

275 CGCATTCCTCGGATAT 258

seq_name: gb_est6:D82221

seq_documentation_block:

LOCUS D82221 375 bp mRNA EST 09-FEB-1996
DEFINITION HUMHBC4626 Human pancreatic Islet Homo sapiens cDNA similar to
HLA-B, mRNA sequence.

ACCESSION D82221

VERSION D82221.1 GI:1183739

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE
AUTHORS      1 (bases 1 to 375)
TITLE        Takeda,J.
JOURNAL      Human pancreatic islet ESTs
COMMENT      Unpublished (1995)
              On Apr 14, 1993 this sequence version replaced gi:785255.
              Contact: Jun Takeda
              Institute for Molecular and Cellular Regulation, Gunma University
              3-39-15 Showa-machi, Maebashi Gunma 371, Japan
              Tel: 272-20-8856
              Fax: 272-20-8896
              Email: jtakeda@sb.gunma-u.ac.jp.

FEATURES
source
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone_lib="Human pancreatic islet"
  /note="Vector: Lambda ZAPII; Site_1: Eco RI; Site_2: Xho
  I; mRNA was prepared from normal adult human islets. cDNA
  was directionally synthesized from the Xho I in the vector
  to the EcoRI site. cDNA was size fractionated to remove
  sequences <1000 bp in size."
BASE COUNT   75 a 124 c 118 g 55 t      3 others
ORIGIN
alignment_scores:
  Quality: 29.00      Length: 6
  Ratio: 4.833       Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-08-653-294-4 x D82221 ..
Align seg 1/1 to: D82221 from: 1 to: 375

1 ArgilleAlaLeuArgTyr 6
|||||
318 CGGATCGCGCTCCGNTAC 335

seq_name: gb_est18:AA710058

seq_documentation_block:
LOCUS      AA710058      390 bp      mRNA      EST      24-DEC-1997
DEFINITION vt39g09.r1 Barstead mouse proximal colon MPLRB6 Mus musculus cDNA
clone IMAGE:1165504 5', mRNA sequence.
ACCESSION  AA710058
VERSION     AA710058.1 GI:2719976
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS    Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE      1 (bases 1 to 390)
JOURNAL    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
COMMENT    Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
              Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
              Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
              Waterston,R.
              The WashU-HMMI Mouse EST Project
              Unpublished (1996)
              On Sep 12, 1996 this sequence version replaced gi:1291974.
              Contact: Maria M/Mouse EST Project
              WashU-HMMI Mouse EST Project
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: mouseest@watson.wustl.edu
              This clone is available royalty-free through LLNL; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              MGI:631416
              Seq primer: -28m13 rev2 ET from Amersham
              High quality sequence stop: 383.

```

```

FEATURES
source
  /organism="Mus musculus"
  /strain="FVB/N"
  /db_xref="taxon:10090"
  /clone_lib="IMAGE:1165504"
  /dev_stage="7 day juvenile"
  /lab_host="DH10B"
  /note="vector: pT73D-Pac (Pharmacia) with a modified
  polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
  was primed with a Not I - oligo(dT) primer [5'
  TGTTACGATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT
  3']; double-stranded cDNA was ligated to Eco RI adaptors
  [AATTCGGATCCTTG], digested with Not I and cloned into the
  Not I and Eco RI sites of the modified pT73 vector.
  Library constructed by Bob Barstead.
BASE COUNT   112 a 95 c 75 g 108 t
ORIGIN
alignment_scores:
  Quality: 29.00      Length: 6
  Ratio: 4.833       Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-08-653-294-4 x AA710058/rev ..
Align seg 1/1 to reverse of: AA710058 from: 1 to: 390

1 ArgilleAlaLeuArgTyr 6
|||||
241 AGGATAGCACTTAGTAC 224

seq_name: gb_est37:AI945988

seq_documentation_block:
LOCUS      AI945988      403 bp      mRNA      EST      17-AUG-1999
DEFINITION bs19b12.y1 Drosophila melanogaster adult testis library Drosophila
melanogaster cDNA clone bs19b12 5', mRNA sequence.
ACCESSION  AI945988
VERSION     AI945988.1 GI:5736386
KEYWORDS   EST.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
REFERENCE  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
AUTHORS    Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
TITLE      Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
JOURNAL    Andrews, J., Bouffard, G. and Oliver, B.
COMMENT    1 (bases 1 to 403)
              Drosophila melanogaster testis expressed sequence tags
              Unpublished (1999)
              On May 18, 1998 this sequence version replaced gi:3138562.
              Contact: Brian Oliver
              Laboratory of Cellular and Developmental Biology
              NIDDK, National Institutes of Health
              6 Center Drive MSC 2715, Bldg 6, Rm B1-13, Bethesda, MD 20892 USA
              Fax: (301) 496 5239
              Email: oliver@helix.nih.gov,
              http://www.niddk.nih.gov/intran/people/boliver.htm
              Tissue isolation and library construction performed at the National
              Institute of Diabetes and Digestive and Kidney Diseases, NIH (see
              http://www.niddk.nih.gov/intran/people/boliver.htm). DNA sequencing
              and analyses performed by National Institutes of Health Intramural
              Sequencing Center (NISC; see http://www.nisc.nih.gov). Chromatogram
              data were analyzed and evaluated for high quality using the ted
              program (Gleeson T and Hillier L, 1991).
              Plate: 19 row: b column: 12
              Seq primer: M13RPI reverse primer (ABI).
              Location/Qualifiers
              1. .403
              /organism="Drosophila melanogaster"

```



```

/strain="y[*] w[67cl]/Y"
/db_xref="taxon:7227"
/clone="bs19b12"
/clone_lib="Drosophila melanogaster adult testis library"
/sex="male"
/dev_stage="1-5 day adult"
/lab_host="SOLR (Stratagene)"
/Note="Organ: testis; Vector: pBluescript SK (Stratagene);
Site:1: EcoRI; Site:2: Xho I; Testes dissected from 1-5
day adult y[*] w[67cl]/Y males raised at 250C. RNA
isolated using Trizol (Life technologies) and a single
round of Poly(A)+ selection using Oligotex (Qiagen). cDNA
library constructed using Stratagene ZAP-cDNA synthesis
kit. Oligo dt-primed, size fractionated -1-6 kb, and
directionally cloned at EcoRI and XhoI in Uni-ZAP XR.
Following a single round of amplification pBluescript SK
phagemids were mass excised. A distribution channel for
clones is being sought, but not currently available.
Requests for clones cannot be honored."
BASE COUNT      90 a 108 c 123 g 82 t
ORIGIN

```

```

alignment_scores:
  Quality: 29.00      Length: 6
  Ratio: 4.833       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```
alignment_block:
US-08-653-294-4 x AI945988 ..

```

```
Align seg 1/1 to: AI945988 from: 1 to: 403
```

```

1 ArgileAlaLeuArgTyr 6
|||||
69 AGGATCGCCTCAGGTAC 86

```

```
seq_name: gb_est20:AA817413
```

```

seq_documentation_block:
LOCUS   AA817413      415 bp      mRNA      EST      25-NOV-1998
DEFINITION   LD22688.5prime LD Drosophila melanogaster embryo pot2 Drosophila
melanogaster cDNA clone LD22688 5prime, mRNA sequence.
ACCESSION   AA817413
VERSION     AA817413.1 GI:2887022
KEYWORDS    EST.
SOURCE      fruit fly.

```

```

ORGANISM    Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 415)
Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
Broksstein,P., Lewis,S. and Rubin,G.M.
BDGP/HMI Drosophila EST Project
Unpublished (1997)
On Jan 19, 1998 this sequence version replaced gi:2045600.
Contact: Harvey D
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: 226 row: H column: 4
High quality sequence stop: 402.
Location/Qualifiers
  1..415
    /organism="Drosophila melanogaster"
    /db_xref="taxon:7227"
    /clone="LD22688"
    /clone_lib="LD Drosophila melanogaster embryo pot2"
    /sex="male and female"
    /dev_stage="0 to 24 hours mixed stage embryonic"

```

```
FEATURES
      source
        1..415

```

```

/lab_host="XLI Blue"
/Note="Organ: embryo; Vector: pot2; Site:1: EcoRI; Site:2:
XhoI; Sized fractionated cDNAs were directly ligated into
pot2."
BASE COUNT      112 a 110 c 103 g 90 t
ORIGIN

```

```

alignment_scores:
  Quality: 29.00      Length: 6
  Ratio: 4.833       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```
alignment_block:
US-08-653-294-4 x AA817413/rev ..

```

```
Align seg 1/1 to reverse of: AA817413 from: 1 to: 415
```

```

1 ArgileAlaLeuArgTyr 6
|||||
403 CGAATAGCGCTCGGTAC 386

```

```
seq_name: gb_est9:AA067065
```

```

seq_documentation_block:
LOCUS   AA067065      430 bp      mRNA      EST      04-FEB-1997
DEFINITION   mm31d10.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
IMAGE:523123 5', mRNA sequence.
ACCESSION   AA067065
VERSION     AA067065.1 GI:1564828
KEYWORDS    EST.
SOURCE      house mouse.

```

```

ORGANISM    Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 430)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
On Sep 12, 1996 this sequence version replaced gi:1393887.
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:316971
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 402.
Location/Qualifiers
  1..430
    /organism="Mus musculus"
    /strain="C57BL/6"
    /db_xref="taxon:10090"
    /clones="IMAGE:523123"
    /clone_lib="Stratagene mouse skin (#937313)"
    /sex="females"
    /tissue_type="whole skin"
    /dev_stage="11 weeks old"
    /lab_host="SOLR (kanamycin resistant)"
    /Note="Organ: skin; Vector: pBluescript SK-; Site:1:
EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 kb; Uni-ZAP XR vector; -5'
adaptor sequence: 5' GAATTCGCACGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

```

```

TITLE
JOURNAL
COMMENT

```

```
FEATURES
      source
        1..430

```

BASE COUNT 118 a 108 c 92 g 112 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-4 x AA067065/rev ..

Align seg 1/1 to reverse of: AA067065 from: 1 to: 430

1 ArgileAlaLeuArgTyr 6
|||||
399 AGGATAGCAGCTTAGGTAC 382

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:29:35 ; Search time 122.56 Seconds
(without alignments)
1.160 Million cell updates/sec

Title: US-08-653-294-5
Perfect score: 29
Sequence: 1 RILLY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : A_Geneseq_36;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|--------|--------------------|
| 1 | 29 | 100.0 | 6 | W47262 | Immunomodulatory p |
| 2 | 29 | 100.0 | 6 | W33782 | Peptide #3 used in |
| 3 | 29 | 100.0 | 10 | R83096 | HLA-B2702 CTL modu |
| 4 | 29 | 100.0 | 10 | W47267 | Immunomodulatory p |
| 5 | 29 | 100.0 | 10 | W47271 | Immunomodulatory p |
| 6 | 29 | 100.0 | 10 | W33789 | Peptide B2702.75-8 |
| 7 | 27 | 93.1 | 10 | W82826 | Cytomodulating lip |
| 8 | 25 | 86.2 | 103 | W54388 | Actinodadura hibi |
| 9 | 25 | 86.2 | 108 | R93159 | Murine monoclonal |
| 10 | 25 | 86.2 | 125 | P80289 | Human Vpre-B prote |
| 11 | 25 | 86.2 | 142 | P80288 | V preB-1 protein. |
| 12 | 25 | 86.2 | 142 | P83001 | V preB-2 protein. |
| 13 | 25 | 86.2 | 365 | R50040 | XMEF2. New nucleic |
| 14 | 24 | 82.8 | 6 | W47261 | Immunomodulatory p |
| 15 | 24 | 82.8 | 6 | W33780 | Peptide #1 used in |
| 16 | 24 | 82.8 | 10 | R41208 | Peptide fragment o |
| 17 | 24 | 82.8 | 10 | R41212 | Peptide fragment o |
| 18 | 24 | 82.8 | 10 | R83062 | HLA-B2702 CTL modu |
| 19 | 24 | 82.8 | 10 | R83075 | HLA-B2702 CTL modu |
| 20 | 24 | 82.8 | 10 | R83094 | HLA-B2702 CTL modu |
| 21 | 24 | 82.8 | 10 | R95413 | Alpha1-helix of HL |
| 22 | 24 | 82.8 | 10 | R95427 | HLA-B2702.75-84(L) |
| 23 | 24 | 82.8 | 10 | R95423 | HLA-B2702.75-84, C |
| 24 | 24 | 82.8 | 10 | R95425 | HLA-B2702.75-84(D) |
| 25 | 24 | 82.8 | 10 | W07524 | T-cell modulating |
| 26 | 24 | 82.8 | 10 | W07512 | T-cell modulating |
| 27 | 24 | 82.8 | 10 | W07513 | T-cell modulating |
| 28 | 24 | 82.8 | 10 | W07514 | T-cell modulating |
| 29 | 24 | 82.8 | 10 | W47265 | Immunomodulatory p |
| 30 | 24 | 82.8 | 10 | W47269 | Immunomodulatory p |
| 31 | 24 | 82.8 | 10 | W33784 | Peptide B2702.75-8 |
| 32 | 24 | 82.8 | 10 | W33785 | Peptide B2705.75-8 |
| 33 | 24 | 82.8 | 10 | W33787 | Peptide B2702.75-8 |
| 34 | 24 | 82.8 | 12 | R95429 | HLA-B2702 84-79-84 |

Peptide B2702.84-7
Immunomodulating d
Glucose transport
HLA-B2702 CTL modu
Peptide B2702.70-8
Human MHC 1 and HL
Human HLA-B27-(62-
Human [Phe74]-HLA-
Human MHC 1 alpha
Human MHC 1 alpha
Human MHC 1 alpha

35 24 82.8 12 1 W33798
36 24 82.8 12 1 W33799
37 24 82.8 13 1 W29421
38 24 82.8 15 1 R92912
39 24 82.8 15 1 W33795
40 24 82.8 17 1 R71440
41 24 82.8 17 1 R71442
42 24 82.8 17 1 R71443
43 24 82.8 17 1 R71425
44 24 82.8 17 1 R71426
45 24 82.8 17 1 R71428

ALIGNMENTS

RESULT 1
W47262
ID W47262 standard; peptide; 6 AA.
AC W47262;
DT 22-MAY-1998 (first entry)
DE Immunomodulatory peptide.
KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
KW transplant rejection; treatment; autoimmune disease.
OS Homo sapiens.
QS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 1..6 /note= "at least one of the amino acids is the D-isomer"
FT PN WD9744052-A1.
PD 27-NOV-1997.
PE 23-APR-1997; U06705
PR 22-MAY-1996; US-651650.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM;
DR WPI; 98-018220/02.
PT Novel immunomodulatory peptide-type compound - useful for inhibiting
transplant rejection
PS Claim 10; Page 36; 41pp; English.
CC The present sequence is an immunomodulatory peptide, which
CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
CC in a pharmaceutical composition together with a subtherapeutic dose
CC of an immunosuppressant, to extend the period of acceptance of a
CC transplant from a major histocompatibility complex (MHC) unmatched
CC donor, i.e. to inhibit transplant rejection. It can also be used in
CC the treatment of autoimmune diseases.
CC Peptides using the D-form amino acids are more effective
CC immunomodulators than their diastereomers or enantiomers.
SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
Db 1 RILLY 6

RESULT 2
W33782
ID W33782 standard; peptide; 6 AA.
AC W33782;
DT 19-JUN-1998 (first entry)
DE Peptide #3 used in immunomodulating dimer peptide.
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
KW rejection.
OS Synthetic.
OS Homo sapiens.
PN WO9744351-A1.

PD 27-NOV-1997.
PE 22-MAY-1997; U08689.
PR 24-MAY-1996; US-653294.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Beulow R, Clayberger C, Krensky AM;
DR WPI: 98-086530/08.
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
PT alpha-1 domain, used for preventing rejection of transplants or
PT treating autoimmune diseases
PS Claim 15; Page 35; 41pp; English.
CC This sequence represents a specifically claimed peptide which forms part
CC of the immunomodulating dimer peptides of the invention. A peptide-type
CC compound or variant is claimed which has immunomodulating activity,
CC including the N-terminal acylated and/or C-terminal amidated or
CC esterified forms of up to 60 amino acids, where the peptide-type compound
CC comprises the formula: A-B, where A, B = (R aa76-77L) (aa79-84) or
CC (aa84-79) (Laa77-76R); aa76 = E or V; aa77 = D, S or N; aa79 = R or G;
CC aa80 = I or N; aa81, aa84 = a hydrophobic or small amino acid; aa82 = R
CC or L; aa83 = G or R; and aa represents amino acid. The sequence in the
CC brackets may optionally be absent or truncated at any peptide type bond
CC within the brackets. The compounds comprise amino acid sequences related
CC to a Class I HLA-B alpha1 domain (positions 79-84). They can be used to
CC inhibit cytotoxic T-lymphocytes (CTL) from undesirably attacking cells in
CC a host or in vitro. They can also be used in combination with antigenic
CC peptides or proteins of interest to activate CTLs. They can also inhibit
CC the proliferation of T cells in response to anti-CD3. The peptide can be
CC used for preventing rejection of transplants or for treating autoimmune
CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
CC The products can also be used for detection and diagnosis.
SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RILLY 6
Db 1 RILLY 6
RESULT 3
R3096
ID R83096 standard; peptide; 10 AA.
AC R83096;
DT 16-MAY-1996 (first entry)
DE HLA-B2702 CTL modulating peptide (B2702.75-84(L)).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW class I MHC; HLA-B2702.
OS Synthetic.
PN W09526979-A1.
PD 12-OCT-1995.
PE 05-APR-1995; U04349.
PR 05-APR-1994; US-222851.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM, Parham P;
DR WPI: 95-358582/46.
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
PS Example 14; Page 34; 80pp; English.
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of
CC class I major histocompatibility complex (MHC) antigens. This sequence
CC corresponds to residues 75-84 of the alpha-1 domain of the class I MHC
CC HLA-B2702. These sequences can be used to extend the period of
CC acceptance by a recipient of a transplant from an MHC unmatched donor.
CC The peptides are administered to a patient in conjunction with
CC a subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.

SQ Sequence 10 AA;
Query Match 100.0%; Score 29; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RILLY 6
Db 5 RILLY 10
RESULT 4
W47267
ID W47267 standard; peptide; 10 AA.
AC W47267;
DT 22-MAY-1998 (first entry)
DE Immunomodulatory peptide.
KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
KW transplant rejection; treatment; autoimmune disease.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 1..10
FT /note= "at least one of the amino acids is the
FT D-isomer
PN W09744052-A1.
PD 27-NOV-1997.
PE 23-APR-1997; U06705.
PR 22-MAY-1996; US-651650.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM;
DR WPI: 98-018220/02.
PT Novel immunomodulatory peptide-type compound - useful for inhibiting
PT transplant rejection
PS Claim 10; Page 36; 41pp; English.
CC The present sequence is an immunomodulatory peptide, which
CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
CC in a pharmaceutical composition together with a subtherapeutic dose
CC of an immunosuppressant, to extend the period of acceptance of a
CC transplant from a major histocompatibility complex (MHC) unmatched
CC donor, i.e. to inhibit transplant rejection. It can also be used in
CC the treatment of autoimmune diseases.
CC Peptides using the D-form amino acids are more effective
CC immunomodulators than their diastereomers or enantiomers.
SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RILLY 6
Db 5 RILLY 10
RESULT 5
W47271
ID W47271 standard; peptide; 10 AA.
AC W47271;
DT 22-MAY-1998 (first entry)
DE Immunomodulatory peptide.
KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
KW transplant rejection; treatment; autoimmune disease.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 1..10
FT /note= "at least one of the amino acids is the
FT D-isomer
PN W03744052-A1.
PD 27-NOV-1997.

PF 23-APR-1997; U06705.
PR 22-MAY-1996; US-651650.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM;
DR WPI; 98-018220/02.
PT Novel immunomodulatory peptide-type compound - useful for inhibiting
transplant rejection
PS Claim 10; Page 36; 41pp; English.
CC The present sequence is an immunomodulatory peptide, which
comprises a Class I HLA-B alpha-1 domain sequence. It can be used
in a pharmaceutical composition together with a subtherapeutic dose
of an immunosuppressant, to extend the period of acceptance of a
transplant from a major histocompatibility complex (MHC) unmatched
donor, i.e. to inhibit transplant rejection. It can also be used in
the treatment of autoimmune diseases.
CC Peptides using the D-form amino acids are more effective
CC immunomodulators than their diastereomers or enantiomers.
SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
DB 5 RILLY 10

RESULT 6

W33789 ID W33789 standard; peptide; 10 AA.
AC W33789;
DE Peptide B2702.75-84L81 tested for immunomodulating activity.
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
KW rejection.
OS Synthetic.
OS Homo sapiens.
PN W09744351-AL.
PD 27-NOV-1997.
PF 22-MAY-1996; U08689.
PR 24-MAY-1996; US-653294.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Beulow R, Clayberger C, Krensky AM;
DR WPI; 98-086530/08.
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
alpha-1 domain, used for preventing rejection of transplants or
treating autoimmune diseases
PS Example 1; Page 19; 41pp; English.
CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
activity. A peptide-type compound or variant is claimed which has
immunomodulating activity, including the N-terminal acylated and/or
C-terminal amidated or esterified forms of up to 60 amino acids, where
the peptide-type compound comprises the formula: A-B, where A, B =
(R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
acid. The sequence in the brackets may optionally be absent or truncated
at any peptide type bond within the brackets. The compounds comprise
amino acid sequences related to a Class I HLA-B alpha-1 domain (positions
79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
undesirably attacking cells in a host or in vitro. They can also be
used in combination with antigenic peptides or proteins of interest to
activate CTLs. They can also inhibit the proliferation of T cells in
response to anti-CD3. The peptide can be used for preventing rejection
of transplants or for treating autoimmune diseases, e.g. diabetes,
rheumatoid arthritis and lupus erythematosus. The products can also be
used for detection and diagnosis.
SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
DB 5 RILLY 10

RESULT 7

W82826 ID W82826 standard; peptide; 10 AA.
AC W82826;
DE 28-JAN-1999 (first entry)
KW Cytomodulating lipophilic oligopeptide (w).
KW cytotoxic; lymphocytic; inhibition; cytokine; immune system; inflammation;
KW B cell; mononuclear phagocyte; septic shock; rheumatoid arthritis;
KW Crohn's disease; colitis; allergy; transplant.
OS Synthetic.
PN W09846633-AL.
PD 22-OCT-1998.
PF 10-APR-1998; U07231.
PR 23-FEB-1998; US-028083.
PA (SANG-) SANGSTAT MEDICAL CORP.
PI Buelow R, Calas B, Grassy G;
DR WPI; 98-594558/50.
PT New lipophilic peptide(s) that inhibit activation of immune system
cells - used for, e.g. production of cytokine(s) and the
PT inflammatory response, and also for modulating haem-containing
PT enzymes
PS Claim 7; Page 37; 48pp; English.
CC W82804 to W82829 are cytomodulating lipophilic oligopeptides. The
oligopeptides are used to inhibit: (i) activity of lymphocytes
(particularly cytotoxic T cells, but also natural killers, B cells and
mononuclear phagocytes); (ii) production of inflammatory cytokines, and
(iii) an inflammatory response in mammals (e.g. in cases of septic shock,
rheumatoid arthritis (RA), Crohn's disease, colitis and allergy). They
are also used for modulating activity of haem-containing enzymes and for
delaying onset of autoimmune disease (specifically insulin-dependent
diabetes mellitus, RA and systemic lupus erythematosus). In all cases
the oligopeptides may be generated from nucleic acids, and treatments
are in vitro or in vivo. A specific application is treatment of organs
or cells for transplantation, or of the recipient of such transplants.
CC Apart from therapeutic use, the oligopeptides can be used to study
mechanisms of T cell (de)activation and to raise antibodies (used to
identify oligopeptides and to raise anti-idiotypic antibodies that are
competitors of the oligopeptides. The oligopeptides are administered by
bolus injection or infusion, typically at 0.1-50 (preferably 1-25) mg/kg.
CC Treatment with the oligopeptides increases the life of transplants.
SQ Sequence 10 AA;

Query Match 93.1%; Score 27; DB 1; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.4;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
DB 1 RILLY 6

RESULT 8

W54388 ID W54388 standard; Protein; 103 AA.
AC W54388;
DE 18-AUG-1998 (first entry)
DE Actinomadura hibiscus polyketide synthase protein 9.
KW Multienzyme; infection; fungi; yeast; gram-positive bacteria; virus;
OS dihydrobenzo(a)naphthacenequinone aglycon; antibiotic; pradimicin.
PN Actinomadura hibiscus.
PN W09811230-AL.

PD 19-MAR-1998.
PF 13-SEP-1996; U14791.
PA 13-SEP-1996; WO-U14791.
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
PI Dairi T, Oki T;
DR WPI; 98-207391/18.
DR N-PSDB; V26609.
PT Actinomadura polyketide synthase genes - useful for preparation of
PT pradiamicin
PS Disclosure; Page 54; 71pp; English.
CC The Actinomadura hibisca polyketide synthase proteins W54380-W84390 form
CC a multienzyme complex. The enzyme is used for the biosynthesis of a
CC dihydrobenzo(a)naphthacenequinone aglycon preferably a pradiamicin which
CC is an antibiotic useful against systemic fungal infections caused by
CC Candida albicans, Aspergillus fumigatus and Cryptococcus neoformans. It
CC is also active against a wide variety of fungi and yeasts, some
CC Gram-positive bacteria and viruses.
SQ Sequence 103 AA;

Query Match 86.2%; Score 25; DB 1; Length 103;
Best Local Similarity 66.7%; Pred. No. 43;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
|:|:|
Db 7 RVLVY 12

RESULT 9
R93159
ID R93159 standard; Protein; 108 AA.
AC R93159;
DT 24-OCT-1996 (first entry)
DE Murine monoclonal antibody K20 kappa chain variable region.
KW Antibody; light chain; kappa; variable region; K20; Integrin; CD29;
KW beta 1 subunit; humanisation; Hu-K20; immunosuppressant;
KW T cell activation; complementarity determining region; CDR.
OS Mus musculus.
FH Key Location/Qualifiers
FT region 1..23
FT /label= FR1
FT /note= "framework region"
FT region 24..34
FT /label= CDR1
FT /note= "complementarity determining region"
FT region 35..49
FT /label= FR2
FT /note= "framework region"
FT region 50..56
FT /label= CDR2
FT /note= "complementarity determining region"
FT region 57..88
FT /label= FR3
FT /note= "framework region"
FT region 89..94
FT /label= CDR3
FT /note= "complementarity determining region"
FT region 95..108
FT /label= J_kappa1
FR2724393-Al.
PD 15-MAR-1996.
PF 12-SEP-1994; 010858.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (PROT-) PROTEINE PERFORMANCE SA.
PI Bernard A, Cervoni MF, Lefranc MP, Margaritte C;
PI Poul MA;
DR WPI; 96-162083/17.
DR N-PSDB; T26849.
PT Humanisation of non-human immunoglobulin variable regions - for
PT prodn. of humanised antibodies, esp. K20, e.g. as an
PT immunosuppressant

PS Example 1; Fig 2A; 39pp.; French.
CC The present sequence is that of the variable region of the kappa
CC light chain from murine monoclonal antibody K20. The antibody
CC recognises the beta 1 subunit (CD29) of integrins and inhibits
CC activation and proliferation of peripheral T cells induced by
CC anti-CD3 antibodies. Monoclonal antibody K20 is a preferred target
CC for humanisation; the humanised version may be useful as an
CC immunosuppressant. In the humanisation process, the complementarity
CC determining regions (CDRs) of a human antibody with framework
CC regions 70-95% homologous to those of K20 were replaced by the K20
CC CDRs.
SQ Sequence 108 AA;

Query Match 86.2%; Score 25; DB 1; Length 108;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
|:|:|
Db 45 RLLRY 50

RESULT 10
P80289
ID P80289 standard; protein; 125 AA.
AC P80289;
DT 07-DEC-1990 (first entry)
DE Human vpre-B protein.
KW Pre-B cells; non-T acute lymphoblast leukaemia.
PN EP-269127-A.
PD 01-JUN-1988.
PF 27-NOV-1987; 117619
PR 27-NOV-1986; GB-028433.
PR 14-JUL-1987; GB-016497.
PR 14-OCT-1987; GB-024100.
PA (HOFF) HOFFMANN-LA ROCHE AG.
PA Bauer SR, Kudo A, Melchers GF, Sakaguchi N;
DR WPI; 88-148947/22.
DR N-PSDB; N80476.
PT Nucleotide sequence selectively expressed in pre-B cells - used
PT in probes for determining non-T acute lymphoblast leukaemia and
PT for prepn. of polypeptide(s)
PS Disclosure; pp; English.
CC This is encoded by the human pre-B gene which differs from the
CC mouse pre-B-2 gene (of N82441) at several posns. The gene is also
CC selectively expressed in pre-B cell lines. Its pattern of express-
CC ion follows that of vpre-B1 and lambda-5 in the mouse.
CC See also N80470-75, N80475 and N82441-42.
SQ Sequence 125 AA;

Query Match 86.2%; Score 25; DB 1; Length 125;
Best Local Similarity 83.3%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RILLY 6
|:|:|
Db 52 RLLRY 57

RESULT 11
P80288
ID P80288 standard; protein; 142 AA.
AC P80288;
DT 07-DEC-1990 (first entry)
DE v pre-B-1 protein.
KW Pre-B cells; non-T acute lymphoblast leukaemia; ss.
PN EP-269127-A.
PD 01-JUN-1988.
PF 27-NOV-1987; 117619.
PR 27-NOV-1986; GB-028433.
PR 14-JUL-1987; GB-016497.

PR 14-OCT-1987; GB-024100.
PA (HOFF) HOFFMANN-LA ROCHE AG.
PI Bauer SR, Kudo A, Melchers GF, Sakaguchi N;
DR WPI: 88-148947/22.
DR N-PSDB; N80475.
PT Nucleotide sequence selectively expressed in pre-B cells - used
PT in probes for determining non-T acute lymphoblast leukaemia and
PT for prepn. of polypeptide(s)
PS Disclosure: pp; English.
CC This is encoded by the genomic form of the variable region pre-B-1
CC sequence. The gene is not rearranged during pre-B cell development
CC and is 4.6 kb upstream of the lambda-5 gene. This protein can
CC associate with itself or with heavy chain V domains expressed in
CC pre-B cells. Vpre-B-1 and lambda-5 form a complete V domain via
CC non-covalent bonds. The gene is expressed only in pre-B cell lines.
CC See also N80470-74, N80476-77 and N82441-42.
SQ Sequence 142 AA;

Query Match 86.2%; Score 25; DB 1; Length 142;
Best Local Similarity 83.3%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RILRY 6
| | | | |
Db 66 RFLRY 71

RESULT 12
ID P83001 standard; protein; 142 AA.
AC P83001;
DT 07-DEC-1990 (first entry)
DE V pre-B-2 protein.
KW pre-B cells; non-T acute lymphoblast leukaemia.
PN EP-269127-A.
PD 01-JUN-1988.
PF 27-NOV-1987; 117619.
PR 27-NOV-1986; GB-028433.
PR 14-JUL-1987; GB-016497.
PR 14-OCT-1987; GB-024100.
PI (HOFF) HOFFMANN-LA ROCHE AG.
PI Bauer SR, Kudo A, Melchers GF, Sakaguchi N;
DR WPI: 88-148947/22.
DR N-PSDB; N82441.
PT Nucleotide sequence selectively expressed in pre-B cells - used
PT in probes for determining non-T acute lymphoblast leukaemia and
PT for prepn. of polypeptide(s)
PS Disclosure: pp; English.
CC This is encoded by the genomic form of the variable region pre-B-2
CC sequence. The gene is not rearranged during pre-B cell development.
CC This protein may associate with itself or with heavy chain V domains
CC expressed in pre-B cells. The gene is expressed only in pre-B cell lines.
CC See also N80470-75, N80476-77 and N82442.
SQ Sequence 142 AA;

Query Match 86.2%; Score 25; DB 1; Length 142;
Best Local Similarity 83.3%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RILRY 6
| | | | |
Db 66 RFLRY 71

RESULT 13
ID R50040 standard; protein; 365 AA.
AC R50040;
DT 22-SEP-1994 (first entry)
DE XMEF2.
KW MEF2; myocyte-specific enhancer factor; isoform;

KW myocyte-specific transcription enhancer factor; muscle; atrophy.
OS Homo sapiens.
PN W09405776-A.
PD 17-MAR-1994.
PF 07-SEP-1993; U08386.
PR 04-SEP-1992; US-939898.
PA (CHIL-) CHILDRENS MEDICAL CENT.
PI Nadal-Ginard B;
DR WPI: 94-101181/12.
DR N-PSDB; Q58741.
PT New nucleic acid encoding MEF2-protein family member - i.e.
PT myocyte-specific transcription enhancing factor, useful to
PT increase muscle mass and to counteract muscle atrophy, e.g. as in
PT muscular dystrophy
PS Claim 6; Page 64-66; 129pp; English.
CC MEF2 transcription factors can be used to produce transgenic animals
CC with increased muscle mass, to prevent or counteract muscle atrophy
CC in humans or animals suffering a pathological muscular condition, or
CC to develop pharmacological agents which regulate the expression of
CC muscle-tissue genes. MEF2 isoforms are given in Q58740-44,
CC R50038-45 and R57772. Nucleic acid encoding a member of the MEF2
CC protein family pref. comprises a sequence encoding at least eleven
CC consecutive glutamine residues or a sequence encoding the amino acid
CC sequence SEEELEL (R57773).
SQ Sequence 365 AA;

Query Match 86.2%; Score 25; DB 1; Length 365;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILRY 6
| | | | |
Db 64 RVLRY 69

RESULT 14
W47261
ID W47261 standard; peptide; 6 AA.
AC W47261;
DT 22-MAY-1998 (first entry)
DE Immunomodulatory peptide.
KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
KW transplant rejection; treatment; autoimmune disease.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 1..6
FT /note= "at least one of the amino acids is the
D-isomer
PN W09744052-A1.
PD 27-NOV-1997.
PF 23-APR-1997; U06705.
PR 22-MAY-1996; US-651650.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM;
DR WPI: 98-018220/02.
PT Novel immunomodulatory peptide-type compound - useful for inhibiting
PT transplant rejection
PS Claim 10; Page 36; 41pp; English.
CC The present sequence is an immunomodulatory peptide, which
CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
CC in a pharmaceutical composition together with a subtherapeutic dose
CC of an immunosuppressant, to extend the period of acceptance of a
CC transplant from a major histocompatibility complex (MHC) unmatched
CC donor, i.e. to inhibit transplant rejection. It can also be used in
CC the treatment of autoimmune diseases.
CC Peptides using the D-form amino acids are more effective
CC immunomodulators than their diastereomers or enantiomers.
SQ Sequence 6 AA;

Query Match 82.8%; Score 24; DB 1; Length 6;

Best Local Similarity 83.3%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RILRY 6
|||
Db 1 RIALRY 6

RESULT 15

W33780
ID W33780 standard; peptide; 6 AA.
AC W33780;
DT 19-JUN-1998 (first entry)
DE Peptide #1 used in immunomodulating dimer peptide.
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
KW rejection.
OS Synthetic.
OS Homo sapiens.
PN W09744351-A1.
PD 27-NOV-1997.
PF 22-MAY-1997; U08689.
PR 24-MAY-1996; US-653294.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Beulow R, Clayberger C, Krensky AM;
DR WPI; 98-086530/08.
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
PT alpha-1 domain, used for preventing rejection of transplants or
PT treating autoimmune diseases
PS Claim 15; Page 35; 4pp; English.
CC This sequence represents a specifically claimed peptide which forms part
CC of the immunomodulating dimer peptides of the invention. A peptide-type
CC compound or variant is claimed which has immunomodulating activity,
CC including the N-terminal acylated and/or C-terminal amidated or
CC esterified forms of up to 60 amino acids, where the peptide-type compound
CC comprises the formula: A-B, where A, B = (R aa76-77L) (aa79-84) or
CC (aa84-79) (Laa77-76R); aa76 - E or V; aa77 - D, S or N; aa79 - R or G;
CC aa80 - I or N; aa81, aa84 - a hydrophobic or small amino acid; aa82 - R
CC or L; aa83 - G or R; and aa represents amino acid. The sequence in the
CC brackets may optionally be absent or truncated at any peptide type bond
CC within the brackets. The compounds comprise amino acid sequences related
CC to a Class I HLA-B alpha domain (positions 79-84). They can be used to
CC inhibit cytotoxic T-lymphocytes (CTL) from undesirably attacking cells in
CC a host or in vitro. They can also be used in combination with antigenic
CC peptides or proteins of interest to activate CTLs. They can also inhibit
CC the proliferation of T cells in response to anti-CD3. The peptide can be
CC used for preventing rejection of transplants or for treating autoimmune
CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
CC The products can also be used for detection and diagnosis.
SQ Sequence 6 AA;

Query Match 82.8%; Score 24; DB 1; Length 6;
Best Local Similarity 83.3%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RILRY 6
|||
Db 1 RIALRY 6

Search completed: February 8, 2000, 01:29:36
Job time: 1748 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 11:54:14 ; Search time 117.7 Seconds
(without alignments)
2.405 Million cell updates/sec

Title: US-08-653-294-5

Perfect score: 29

Sequence: 1 RILLRY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

PIR62.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 29 | 100.0 | 133 | 2 D72110 | hypothetical prote |
| 2 | 29 | 100.0 | 458 | 2 A71467 | probable phosphol |
| 3 | 29 | 100.0 | 458 | 2 D72012 | phosphoglucutase |
| 4 | 29 | 100.0 | 532 | 2 G70986 | probable coA ligas |
| 5 | 28 | 96.6 | 616 | 2 JC4084 | prolyl endopeptida |
| 6 | 27 | 93.1 | 332 | 2 B71508 | hypothetical prote |
| 7 | 26 | 89.7 | 298 | 2 S53849 | ribosomal protein |
| 8 | 26 | 89.7 | 383 | 2 T15698 | hypothetical prote |
| 9 | 26 | 88.7 | 497 | 1 B42902 | methylnalonate-sem |
| 10 | 26 | 89.7 | 502 | 2 D70806 | probable fadbl7 pr |
| 11 | 26 | 89.7 | 617 | 2 C72670 | hypothetical prote |
| 12 | 26 | 89.7 | 1203 | 2 A33165 | finger protein sdc |
| 13 | 26 | 89.7 | 3712 | 2 S18253 | laminin alpha-1 ch |
| 14 | 25 | 86.2 | 103 | 2 JC5858 | polyketide synthas |
| 15 | 25 | 86.2 | 132 | 2 C70161 | ribosomal protein |
| 16 | 25 | 86.2 | 137 | 2 A72666 | hypothetical prote |
| 17 | 25 | 86.2 | 139 | 2 S00258 | VpreB protein - hu |
| 18 | 25 | 86.2 | 142 | 2 B28344 | VpreB protein prec |
| 19 | 25 | 86.2 | 142 | 2 A28344 | VpreB protein prec |
| 20 | 25 | 86.2 | 145 | 2 I57832 | Vpre-B protein - h |
| 21 | 25 | 86.2 | 145 | 2 T13342 | hypothetical prote |
| 22 | 25 | 86.2 | 217 | 2 A42644 | neuE protein - Esc |
| 23 | 25 | 86.2 | 229 | 2 G71322 | hypothetical prote |
| 24 | 25 | 86.2 | 242 | 2 JC5883 | myocyte enhancer f |
| 25 | 25 | 86.2 | 339 | 2 JC5882 | myocyte enhancer f |
| 26 | 25 | 86.2 | 349 | 2 JC5881 | myocyte enhancer f |
| 27 | 25 | 86.2 | 349 | 2 JC4221 | transcription enha |
| 28 | 25 | 86.2 | 365 | 2 A39481 | serum response fac |
| 29 | 25 | 86.2 | 372 | 2 S51137 | ATP-dependent 26S |
| 30 | 25 | 86.2 | 500 | 2 D64889 | probable phenylace |

| | | | | | |
|----|----|------|------|----------|--------------------|
| 31 | 25 | 86.2 | 549 | 2 T03983 | rf2 nuclear restor |
| 32 | 25 | 86.2 | 577 | 2 E71364 | hypothetical prote |
| 33 | 25 | 86.2 | 735 | 2 T13646 | hypothetical prote |
| 34 | 25 | 86.2 | 817 | 1 RRVGCT | RNA-directed RNA p |
| 35 | 25 | 86.2 | 817 | 1 RRVGCT | RNA-directed RNA p |
| 36 | 25 | 86.2 | 817 | 1 RRVGCR | RNA-directed RNA p |
| 37 | 25 | 86.2 | 818 | 2 S05456 | hypothetical prote |
| 38 | 25 | 86.2 | 830 | 2 S25198 | vacuolar membrane |
| 39 | 25 | 86.2 | 832 | 2 S19418 | probable membrane |
| 40 | 25 | 86.2 | 1412 | 2 T01610 | RNA-directed DNA p |
| 41 | 25 | 86.2 | 1612 | 2 S59969 | DNA topoisomerase |
| 42 | 25 | 86.2 | 1626 | 2 A39242 | DNA topoisomerase |
| 43 | 24 | 82.8 | 27 | 4 I52725 | hypothetical BCL2/ |
| 44 | 24 | 82.8 | 73 | 2 S38695 | class II histocomp |
| 45 | 24 | 82.8 | 76 | 2 I68913 | MHC protein - cott |

ALIGNMENTS

RESULT 1

D72110

hypothetical protein - Chlamydia pneumoniae (strain CWL029)

C:Species: Chlamydia pneumoniae

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Apr-1999

C:Accession: D72110

R:Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,

Nature Genet. 21, 385-389, 1999

A>Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606

A:Accession: D72110

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-133 <ARN>

A:Cross-references: GB:AE001604; GB:AE001363; NID:g4376438; PID:g4376450

A:Experimental source: strain CWL029

C:Genetics:

A:Gene: CP0181

Query Match

Best Local Similarity 100.0%; Score 29; DB 2; Length 133;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLRY 6

DB 95 RILLRY 100

RESULT 2

A71467

probable phosphoglucutase - Chlamydia trachomatis (serotype D, strain UW3/Cx)

C:Species: Chlamydia trachomatis

C>Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 29-Sep-1999

C:Accession: A71467

R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche

Science 282, 754-759, 1998

A>Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t

A:Reference number: A71570; MUID:99000809

A:Accession: A71467

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-458 <ARN>

A:Cross-references: GB:AE001354; GB:AE001273; NID:g3329280; PIDN:AAC68412.1; PID:g332

A:Experimental source: serotype D, strain UW-3/Cx

C:Genetics:

A:Gene: mrsa

C:Superfamily: probable phosphorylating protein ureC

Query Match

Best Local Similarity 100.0%; Score 29; DB 2; Length 458;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

C;Date: 12-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 09-Sep-1997
C;Accession: JC4084
R;Robinson, K.A.; Bartley, D.A.; Robb, F.T.; Schreier, H.J.
Gene 152, 103-106, 1995
A;Title: A gene from the hyperthermophile Pyrococcus furiosus whose deduced product 1
A;Reference number: JC4084; MUID:95129900
A;Accession: JC4084
A;Molecule type: DNA
A;Residues: 1-616 <ROB>
A;Cross-references: GB:008343; NID:9475591; PID:9475592
C;Keywords: hydrolase; oligopeptidase
F;477,561,592/Active site: Ser, Asp, His #status predicted

Query Match 96.6%; Score 28; DB 2; Length 616;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
Db 334 RVLRY 339
|:|||||

RESULT 6
B71508
hypothetical protein CT484 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 21-Nov-1998
C;Accession: B71508
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche
Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t
A;Reference number: A71570; MUID:99000809
A;Accession: B71508
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-332 <ARN>
A;Cross-references: GB:AE001322; GB:AE001273; NID:93328916; PID:93328920
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: CT484

Query Match 93.1%; Score 27; DB 2; Length 332;
Best Local Similarity 83.3%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
Db 19 RLLRY 24
|:|||||

RESULT 7
S53849
ribosomal protein S3 - Acanthamoeba castellanii mitochondrion (SGC6)
C;Species: mitochondrion Acanthamoeba castellanii
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 26-Feb-1999
C;Accession: S53849
R;Burger, G.; Plante, I.; Lonergan, K.M.; Gray, M.W.
J. Mol. Biol. 245, 522-537, 1995
A;Title: The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba castellanii: c
A;Reference number: S53825; MUID:95147275
A;Accession: S53849
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-298 <EUR>
A;Cross-references: GB:U12386; NID:9562028; PID:9562053
A;Experimental source: strain Neff; ATCC 30010
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC6
C;Keywords: mitochondrion

```

Query Match 89.7%; Score 26; DB 2; Length 298;
Best Local Similarity 66.7%; Pred. No. 60;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILRY 6
|:|:|
DB 132 RVILRY 137

RESULT 8
ti5698
hypothetical protein C29F5.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: TI5698

R:Miller, N.

submitted to the EMBL Data Library, July 1995

A:Description: The sequence of C. elegans cosmid C29F5.

A:Reference number: 218390

A:Accession: TI5698

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-383 <MIL>

A:Cross-references: EMBL:U23524; NID:G746562; PID:G746563; PIDN:AAC46820.1; CESP:C29F5.1

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:C29F5.1

A:Introns: 28/3; 86/3; 154/2; 219/2; 281/3

Query Match 89.7%; Score 26; DB 2; Length 383;
Best Local Similarity 83.3%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILRY 6
|:|:|
DB 200 RILRY 205

RESULT 9
B42902

methylenaldehyde-semialdehyde dehydrogenase (acylating) (EC 1.2.1.27) - Pseudomonas aeruginosa

C:Species: Pseudomonas aeruginosa

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: B42902; S27602

R:Steele, M.I.; Lorenz, D.; Hatter, K.; Park, A.; Sokatch, J.R.

J. Biol. Chem. 267, 13585-13592, 1992

A:Title: Characterization of the mmsAB operon of Pseudomonas aeruginosa PAO encoding methylenaldehyde-semialdehyde dehydrogenase

A:Reference number: A42902; MUID:92317087

A:Accession: B42902

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-497 <STE>

A:Cross-references: EMBL:M84911; NID:G151360; PIDN:AAA25891.1; PID:G151362

A:Note: sequence extracted from NCB1 backbone (NCBIP:107707)

C:Genetics:

A:Gene: mmsA

C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

C:Keywords: Oxidoreductase

F:43-301/Domain: aldehyde dehydrogenase homology <ALDD>

F:282/Active site: Cys #status predicted

Query Match 89.7%; Score 26; DB 1; Length 497;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILRY 6
|:|:|
DB 68 RVILRY 73

RESULT 10

D70806

Probable fadD17 protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 17-Jul-1998

C:Accession: D70806

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987

A:Accession: D70806

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-502 <COL>

A:Cross-references: GB:AL022022; GB:AL123456; NID:G3261554; PID:el254634; PID:G292444

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: fadD17

Query Match 89.7%; Score 26; DB 2; Length 502;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILRY 6
|:|:|
DB 406 RVLMRY 411

RESULT 11

C72670

hypothetical protein APE0785 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: C72670

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A:Reference number: A72450; MUID:99310339

A:Accession: C72670

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-617 <KAW>

A:Cross-references: DDBJ:AP000060; NID:G5104188; PIDN:BAA79763.1; PID:G1043549; PID:G

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0785

Query Match 89.7%; Score 26; DB 2; Length 617;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILRY 6
|:|:|
DB 330 RVLMRY 335

RESULT 12

A33165

finger protein sdc-1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 09-Sep-1997

C:Accession: S15093; A33165

R:Nonet, M.L.; Meyer, B.J.

Nature 351, 65-68, 1991

A:Title: Early aspects of Caenorhabditis elegans sex determination and dosage compens

A:Reference number: S15093; MUID:91226537

A:Accession: S15093
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1203 <NON>
A:Cross-references: EMBL:X58520; NID:g6848; PID:g6849
C:Genetics:
A:Gene: sdc-1
C:Keywords: DNA binding; nucleus; transcription regulation; zinc finger

Query Match 89.7%; Score 26; DB 2; Length 1203;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
|:|:|
Db 437 RVILY 442

RESULT 13
S18253
laminin alpha-1 chain precursor - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C:Date: 16-Sep-1992 #sequence_revision 24-Jul-1997 #text_change 13-Aug-1999
C:Accession: S28399; S18253
R:Kusche-Gullberg, M.; Garrison, K.; Mackrell, A.J.; Fessler, J.H.
EMBO J. 11, 4519-4527, 1992
A:Title: Laminin A chain: expression during *Drosophila* development and genomic sequence.
A:Reference number: S28399; MUID:93049203
A:Accession: S28399
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-3712 <KUS>
A:Cross-references: GB:M96388; NID:gl57799; PIDN:AAA28662.1; PID:gl57800
R:Garrison, K.; Mackrell, A.J.; Fessler, J.H.
J. Biol. Chem. 266, 22899-22904, 1991
A:Title: *Drosophila* laminin A chain sequence, interspecies comparison, and domain structure.
A:Reference number: S18253; MUID:92078147
A:Accession: S18253
A:Molecule type: mRNA
A:Residues: 1762-3712 <GAR>
A:Cross-references: EMBL:M75882; NID:gl57797; PIDN:AAA28661.1; PID:gl57798
C:Genetics:
A:Gene: FlyBase: LANA
A:Cross-references: FlyBase:FBgn0002526
C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like h
C:Keywords: basement membrane; cell binding; coiled coil; disulfide bond; extracellular
F:273-330/Domain: laminin-type EGF-like homology <LEG>
F:333-400/Domain: laminin-type EGF-like homology <LEG2>
F:541-584/Domain: laminin-type EGF-like homology <LEG1>
F:1778-2115/Domain: III <DOM3>
F:1778-1806/Domain: laminin-type EGF-like homology #status atypical <LE1>
F:1809-1856/Domain: laminin-type EGF-like homology <LE2>
F:1859-1914/Domain: laminin-type EGF-like homology <LE3>
F:1917-1967/Domain: laminin-type EGF-like homology <LE4>
F:1970-2014/Domain: laminin-type EGF-like homology <LE5>
F:2017-2061/Domain: laminin-type EGF-like homology <LE6>
F:2064-2109/Domain: laminin-type EGF-like homology <LE7>
F:2110-2697/Domain: laminin-type EGF-like homology <LE7>
F:2698-3712/Domain: I/II, heptad repeats <DOM2>
F:2698-2863/Domain: repeat G1 <RG1>
F:2864-3048/Domain: repeat G2 <RG2>
F:3049-3223/Domain: repeat G3 <RG3>
F:3079-3200/Domain: laminin G repeat homology <LG3>
F:3334-3528/Domain: repeat G4 <RG4>
F:3529-3712/Domain: repeat G5 <RG5>
F:1847, 1850, 1943, 2024, 2196, 2215, 2267, 2301, 2323, 2482, 2524, 2538, 2569, 2699, 2720, 2890, 2938, 3

Query Match 89.7%; Score 26; DB 2; Length 3712;
Best Local Similarity 83.3%; Pred. No. 7.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
|:|:|
Db 885 RVILY 890

RESULT 14
JC5858
polyketide synthase (EC 2.-.-.-) chain 9 - *Actinomadura hibisca*
C:Species: *Actinomadura hibisca*
C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 17-Mar-1999
C:Accession: JC5858
R:Dairi, T.; Hamano, Y.; Igarashi, Y.; Furumai, T.; Oki, T.
Biosci. Biotechnol. Biochem. 61, 1445-1453, 1997
A:Title: Cloning and nucleotide sequence of the putative polyketide synthase genes fo
A:Reference number: JC5850; MUID:97480928
A:Accession: JC5858
A:Molecule type: DNA
A:Residues: 1-103 <DAI>
A:Cross-references: DDBJ:D87924
C:Comment: This enzyme catalyzes repeated condensation cycles of acyl-CoA, affecting
C:Genetics:
A:Gene: pms9
C:Keywords: transferase

Query Match 86.2%; Score 25; DB 2; Length 103;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
|:|:|
Db 7 RVILY 12

RESULT 15
C70161
ribosomal protein S8 (rpsh) - Lyme disease spirochete
C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 13-Aug-1999
C:Accession: C70161
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kevlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*.
A:Reference number: A70100; MUID:98065943
A:Accession: C70161
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-132 <KLE>
A:Cross-references: GB:AE0001152; GB:AE000783; NID:g2688387; PIDN:AAC66850.1; PID:g268
A:Experimental source: strain B31
C:Superfamily: *Escherichia coli* ribosomal protein S8

Query Match 86.2%; Score 25; DB 2; Length 132;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
|:|:|
Db 61 RVILY 66

Search completed: February 7, 2000, 11:54:16
Job time: 24326 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 00:59:45 ; Search time 63.71 seconds
(without alignments)
2.813 Million cell updates/sec

Title: US-08-653-294-5

Perfect score: 29

Sequence: 1 RILLY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_38.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|---------------|--------------------|
| 1 | 26 | 89.7 | 225 | 1 CBIQ_SALTY | Q05598 salmonella |
| 2 | 26 | 89.7 | 298 | 1 RT03_ACACA | P48754 acanthamoeb |
| 3 | 26 | 89.7 | 496 | 1 MMSA_PSEAE | P28810 pseudomonas |
| 4 | 26 | 89.7 | 1203 | 1 SDC1_CAEEL | P24349 caenorhabdi |
| 5 | 26 | 89.7 | 2569 | 1 LMA3_MOUSE | Q61789 mus musculu |
| 6 | 26 | 89.7 | 3712 | 1 LMA_DROME | Q00174 drosophila |
| 7 | 25 | 86.2 | 132 | 1 RS8_BORBU | O51445 borrelia bu |
| 8 | 25 | 86.2 | 139 | 1 VPRI_MOUSE | P12018 homo sapien |
| 9 | 25 | 86.2 | 142 | 1 VPRI_MOUSE | P13372 mus musculu |
| 10 | 25 | 86.2 | 142 | 1 VPR2_MOUSE | P13373 mus musculu |
| 11 | 25 | 86.2 | 365 | 1 MEFB_HUMAN | Q02080 homo sapien |
| 12 | 25 | 86.2 | 372 | 1 PRS2_METTM | P42811 methanobact |
| 13 | 25 | 86.2 | 415 | 1 TRSA_STRAM | Q07197 streptomyce |
| 14 | 25 | 86.2 | 499 | 1 FEAB_ECOLI | P80668 escherichia |
| 15 | 25 | 86.2 | 515 | 1 MEF2_DROME | P40791 drosophila |
| 16 | 25 | 86.2 | 577 | 1 Y110_TREPA | O83148 treponema p |
| 17 | 25 | 86.2 | 817 | 1 RPPO_CNV | P15187 cucumber ne |
| 18 | 25 | 86.2 | 817 | 1 RPPO_CNV | P17459 cymbidium r |
| 19 | 25 | 86.2 | 817 | 1 RPPO_TBVC | P15962 tomato bush |
| 20 | 25 | 86.2 | 830 | 1 HMT1_SCHPO | Q03592 schizosacch |
| 21 | 25 | 86.2 | 832 | 1 YCZ6_YEAST | P25611 saccharomyc |
| 22 | 25 | 86.2 | 1612 | 1 TP2B_CRILLO | Q64399 cricetus |
| 23 | 25 | 86.2 | 1612 | 1 TP2B_MOUSE | Q64511 mus musculu |
| 24 | 25 | 86.2 | 1626 | 1 TP2B_HUMAN | Q02880 homo sapien |
| 25 | 25 | 86.2 | 1627 | 1 TP2B_CHICK | Q42131 gallus gall |
| 26 | 24 | 82.8 | 101 | 1 Y941_HAEIN | P44082 haemophilus |
| 27 | 24 | 82.8 | 133 | 1 RS8_CHLTR | P28544 chlamydia t |
| 28 | 24 | 82.8 | 147 | 1 SPAM_SALTY | P40612 salmonella |
| 29 | 24 | 82.8 | 168 | 1 Y2DC_ECOLI | P09183 escherichia |
| 30 | 24 | 82.8 | 178 | 1 Y4YJ_RHIFR | P72273 rhizobium f |
| 31 | 24 | 82.8 | 178 | 1 Y4YJ_RHISN | P55718 rhizobium s |
| 32 | 24 | 82.8 | 211 | 1 Y290_METJA | Q57738 methanococc |
| 33 | 24 | 82.8 | 217 | 1 TER4_ECOLI | P09164 escherichia |
| 34 | 24 | 82.8 | 217 | 1 YUJ7_YEAST | P40857 saccharomyc |

| | | | | | |
|----|----|------|-----|--------------|--------------------|
| 35 | 24 | 82.8 | 218 | 1 TER8_PASPI | P51562 pasteurella |
| 36 | 24 | 82.8 | 249 | 1 YE20_ARCFU | Q28852 archaeoglob |
| 37 | 24 | 82.8 | 328 | 1 HAIQ_MOUSE | P14428 mus musculu |
| 38 | 24 | 82.8 | 334 | 1 SRB7_CAEEL | P34142 caenorhabdi |
| 39 | 24 | 82.8 | 338 | 1 IB20_HUMAN | P30467 homo sapien |
| 40 | 24 | 82.8 | 338 | 1 GALE_NEIGO | Q05026 neisseria g |
| 41 | 24 | 82.8 | 339 | 1 GALE_NEIME | Q05624 neisseria m |
| 42 | 24 | 82.8 | 339 | 1 Y83C_METJA | P81324 methanococc |
| 43 | 24 | 82.8 | 345 | 1 PB12_TRYEB | P23734 trypanosoma |
| 44 | 24 | 82.8 | 348 | 1 HIAF_MACMU | P33617 macaca mula |
| 45 | 24 | 82.8 | 359 | 1 IB01_PANTR | P13750 pan troglod |

ALIGNMENTS

RESULT 1
CBIQ_SALTY
ID CBIQ_SALTY STANDARD; PRT; 225 AA.
AC Q05598;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE COBALT TRANSPORT PROTEIN CBIQ.
GN CBIQ.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE: 93273696.
RA ROTH J.R., LAWRENCE J.G., RUBENFIELD M., KIEFFER-HIGGINS S.,
RA CHURCH G.M.;
RT "Characterization of the cobalamin (vitamin B12) biosynthetic genes
of Salmonella typhimurium.";
RL J. Bacteriol. 175:3303-3316(1993).
CC -1- FUNCTION: MAY BE INVOLVED WITH COBALT TRANSPORT IN ASSOCIATION
WITH COBALAMIN BIOSYNTHESIS.
CC -1- PATHWAY: COBALAMIN BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).
CC EMBL: L12006; AAA27266.1; -
DR STYGENE; SG10049; CBIQ.
KW Cobalamin biosynthesis; Transport; Cobalt transport; Transmembrane.
FT TRANSMEM 22 42 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 90 110 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 150 170 POTENTIAL.
SQ SEQUENCE 225 AA; 25983 MW; E926FA4B CRC32;

Query Match 89.7%; Score 26; DB 1; Length 225;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RILLY 6
|:|:|
Db 202 RVLMEY 207

RESULT 2
RT03_ACACA
ID RT03_ACACA STANDARD; PRT; 298 AA.

```

CC CC -1- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensese@isb-sib.ch).
CC -----
CC EMBL; M84911; AAA25891.1; -
CC PIR; B42902; B42902.
CC PIR; S27602; S27602.
CC HSP; P56533; LBW.
CC PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
CC DR DR
CC PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; FALSE_NEG.
CC DR DR
CC PFAM; PF00171; aldehyd; 1.
CC Oxidoreductase; NAD.
CC INIT MET 0
CC ACT_SITE 281 281 BY SIMILARITY.
CC SEQUENCE 496 AA; 53532 MW; 4D33C45B CRC32;
CC -----
CC Query Match 89.7%; Score 26; DB 1; Length 496;
CC Best Local Similarity 66.7%; Pred. No. 52;
CC Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
CC QY 1 RILLRY 6
CC |::|||
CC DB 67 RVMLRY 72
CC -----
CC RESULT 4
CC SDCL CAEEL STANDARD; PRT; 1203 AA.
CC ID SDCL CAEEL STANDARD; PRT; 1203 AA.
CC AC P24349; Q20672.;
CC 01-MAR-1992 (Rel. 21, Created)
CC DT 01-MAR-1992 (Rel. 21, Last sequence update)
CC DT 01-MAR-1992 (Rel. 21, Last sequence update)
CC DT 01-NOV-1997 (Rel. 35, Last annotation update)
CC DE ZINC FINGER PROTEIN SDC-1.
CC DD SDC-1 OR EGL-16 OR F52E10.1.
CC GN Caenorhabditis elegans.
CC OS Caenorhabditis elegans.
CC OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
CC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-BRISTOL N2;
CC RX MEDLINE; 91226537.
CC RT NONET M.L., MEYER B.J.;
CC RA "Early aspects of Caenorhabditis elegans sex determination and dosage
CC RL compensation are regulated by a zinc-finger protein.";
CC RL Nature 351:65-68(1991).
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-BRISTOL N2;
CC RX LLOYD C.;
CC RA Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC CC -1- FUNCTION: EMBRYONIC TRANSCRIPTION FACTOR REGULATING DOWNSTREAM
CC GENES INVOLVED SPECIFICALLY IN THE SEX DETERMINATION AND DOSAGE
CC COMPENSATION PATHWAYS, OR REGULATING OTHER GENES INVOLVED IN THE
CC COORDINATE CONTROL OF BOTH PROCESSES.
CC CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensese@isb-sib.ch).
CC -----
CC EMBL; X58520; CAA41410.1; -
CC EMBL; 254282; CAA91056.1; -

```

DR PIR: A33165; A33165.
 DR PIR: S15093; S15093.
 DR TRANSFAC: T00742; -
 DR WORKPSP: F52E10.1; CE03393.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 2.
 DR PFAM; PF00096; zf-C2H2; 4.
 KW Sexual differentiation; Transcription regulation; Zinc-finger;
 KW Metal-binding; DNA-binding; Developmental protein; Nuclear protein.
 FT DOMAIN 3 15
 FT ZN_FING 117 139
 FT ZN_FING 145 168
 FT ZN_FING 235 256
 FT ZN_FING 270 292
 FT ZN_FING 488 515
 FT ZN_FING 523 545
 FT ZN_FING 654 676
 FT ZN_FING 16 16
 FT CONFLICT 944 945 FS -> D (IN REF. 2).
 FT CONFLICT 944 945 FS -> KT (IN REF. 2).
 SQ SEQUENCE 1203 AA; 139286 MW; 8A760914 CRC32;

Query Match 89.78; Score 26; DB 1; Length 1203;
 Best Local Similarity 66.78; Pred. No. 1.3e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RILLY 6
 |::|||
 Db 437 RVILRY 442

RESULT 5
 LMA3_MOUSE STANDARD: PRT; 2569 AA.
 AC Q61789; Q61788; Q61966;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE LAMININ ALPHA-3 CHAIN PRECURSOR (FRAGMENT).
 GN LAMA3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE=LUNG;
 RX MEDLINE; 95394948;
 RA GALLIANO M.-F., ABERDAM D., AGUZZI A., ORTONNE J.-P., MENEGUZZI G.;
 RT "Cloning and complete primary structure of the mouse laminin alpha 3
 chain. Distinct expression pattern of the laminin alpha 3A and alpha
 3B chain isoforms.";
 RL J. Biol. Chem. 270:21820-21826(1995).
 RN [2]
 RN REVISIONS.
 RA ABERDAM D.;
 RN Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 RC TISSUE=LUNG;
 RX MEDLINE; 94281750.
 RA ABERDAM D., GALLIANO M.-F., MATTEI M.-G., PISANI-SPADAFORA A.,
 RA ORTONNE J.-P., MENEGUZZI G.;
 RT "Assignment of mouse nectin genes to chromosomes 1 and 18.";
 RL Mamm. Genome 5:229-233(1994).
 RN [4]
 RN SEQUENCE OF 1052-1770 FROM N.A.
 RC TISSUE=LUNG;
 RX MEDLINE; 94363405.
 RA ABERDAM D., AGUZZI A., BAUDOUIN C., GALLIANO M.-F., ORTONNE J.-P.,
 RA MENEGUZZI G.;
 RT "Developmental expression of nectin adhesion protein (laminin-5)
 subunits suggests multiple morphogenic roles.";
 RL Cell Adhes. Commun. 2:115-129(1994).
 CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ

IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
 CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
 WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
 -1- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION
 VIA INTEGRIN ALPHA-3/BETA-1 IN FOCAL ADHESION AND INTEGRIN ALPHA-
 6/BETA-4 IN HEMIDESMOSOMES, (2) SIGNAL TRANSDUCTION VIA TYROSINE
 PHOSPHORYLATION OF P125-FAK AND P80, (3) DIFFERENTIATION OF
 KERATINOCYTES (BY SIMILARITY).
 -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
 DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
 TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
 COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
 THE ALPHA-3 CHAIN IS A SUBUNIT OF LAMININ-5 (EPILIGRIN/KALININ/
 NICEIN). AND POSSIBLY ALSO A COMPONENT OF LAMININ-6 (K-LAMININ)
 AND LAMININ-7 (KS-LAMININ).
 -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
 MEMBRANES (MAJOR COMPONENT).
 -1- ALTERNATIVE PRODUCTS: THE TWO ISOFORMS A AND B, WHICH DIFFER
 IN THEIR N-TERMINAL ARE DERIVED BY ALTERNATIVE SPLICING OF THE
 SAME GENE. THE SEQUENCE SHOWN HERE IS THAT OF THE LARGER ISOFORM
 B.
 -1- TISSUE SPECIFICITY: BASAL MEMBRANE OF THE UPPER ALIMENTARY TRACT
 AND URINARY AND NASAL EPITHELIA, SALIVARY GLANDS AND TEETH (BOTH
 VARIANTS). ISOFORM A IS PREDOMINANTLY EXPRESSED IN SKIN, HAIR
 FOLLICLES AND DEVELOPING NEURONS OF THE TRIGEMINAL GANGLION.
 ISOFORM B WAS FOUND IN BRONCHI, ALVEOLI, STOMACH, INTESTINAL
 CRYPTS, WHISKER PADS, CNS, TELECEPHALIC NEUROECTODERM, THALAMUS,
 RATHKE'S POUCH, AND PERIVENTRICULAR SUBENDYMAL GERMINAL LAYER.
 -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 -1- DOMAIN: DOMAINS IV AND G ARE GLOBULAR.
 -1- SIMILARITY: CONTAINS 1.5 LAMININ EGF-LIKE DOMAINS.
 -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV (DOMAIN IV' IS NOT
 SIMILAR TO LAMININ DOMAIN IV).
 -1- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; X84014; CAA58837.1; -
 CC EMBL; X84013; CAA58836.1; -
 CC EMBL; L20478; AAA68091.1; -
 CC HSSP; P02468; ITLE.
 CC MGD; MGI:99909; LAMA3.
 CC PROSITE; PS00022; EGF_1; 4.
 CC PROSITE; PS01186; EGF_2; 1.
 CC PROSITE; PS01248; LAMININ_TYPE_EGF; 4.
 CC PFAM; PF00052; laminin_B; 1.
 CC PFAM; PF00053; laminin_EGF; 4.
 CC PFAM; PF00054; laminin_G; 2.
 CC KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal;
 KW Alternative splicing.
 FT NONTER 1 1
 FT SIGNAL <1 28 POTENTIAL.
 FT CHAIN 29 2569 LAMININ ALPHA-3 CHAIN.
 FT DOMAIN 29 498 DOMAIN IV'.
 FT DOMAIN 499 700 DOMAIN III B.
 FT DOMAIN 546 700 3.5 X LAMININ EGF-LIKE REPEATS.
 FT DOMAIN 546 589 LAMININ EGF-LIKE 1.
 FT DOMAIN 590 639 LAMININ EGF-LIKE 2.
 FT DOMAIN 640 690 LAMININ EGF-LIKE 3.
 FT DOMAIN 691 700 LAMININ EGF-LIKE 4 (N-TERMINAL).
 FT DOMAIN 701 889 LAMININ DOMAIN IV (DOMAIN IV A).
 FT DOMAIN 890 1057 3 X LAMININ EGF-LIKE REPEATS (DOMAIN III
 A).
 FT DOMAIN 890 922 LAMININ EGF-LIKE 4 (C-TERMINAL).
 FT DOMAIN 923 969 LAMININ EGF-LIKE 5.

FT DOMAIN 970 1022 LAMININ EGF-LIKE 6.
 FT DOMAIN 1023 1057 LAMININ EGF-LIKE 7 (INCOMPLETE).
 FT DOMAIN 1058 1648 DOMAIN II AND I (HEPAT REPEATS).
 FT DOMAIN 1649 2369 5 X LAMININ G-LIKE REPEATS (DOMAIN G).
 FT DOMAIN 1649 1825 LAMININ G-LIKE 1.
 FT DOMAIN 1826 1994 LAMININ G-LIKE 2.
 FT DOMAIN 1995 2209 LAMININ G-LIKE 3.
 FT DOMAIN 2210 2385 LAMININ G-LIKE 4.
 FT DOMAIN 2386 2569 LAMININ G-LIKE 5.
 FT DOMAIN 1090 1219 COILED COIL (POTENTIAL).
 FT DOMAIN 1251 1296 COILED COIL (POTENTIAL).
 FT DOMAIN 1327 1404 COILED COIL (POTENTIAL).
 FT DOMAIN 1450 1477 COILED COIL (POTENTIAL).
 FT DOMAIN 1557 1622 COILED COIL (POTENTIAL).
 FT SITE 1513 1515 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 546 553 BY SIMILARITY.
 FT DISULFID 548 560 BY SIMILARITY.
 FT DISULFID 562 571 BY SIMILARITY.
 FT DISULFID 574 587 BY SIMILARITY.
 FT DISULFID 590 605 BY SIMILARITY.
 FT DISULFID 592 612 BY SIMILARITY.
 FT DISULFID 614 623 BY SIMILARITY.
 FT DISULFID 626 637 BY SIMILARITY.
 FT DISULFID 640 652 BY SIMILARITY.
 FT DISULFID 642 659 BY SIMILARITY.
 FT DISULFID 661 670 BY SIMILARITY.
 FT DISULFID 673 688 BY SIMILARITY.
 FT DISULFID 923 932 BY SIMILARITY.
 FT DISULFID 925 939 BY SIMILARITY.
 FT DISULFID 942 951 BY SIMILARITY.
 FT DISULFID 954 967 BY SIMILARITY.
 FT DISULFID 970 982 BY SIMILARITY.
 FT DISULFID 972 981 BY SIMILARITY.
 FT DISULFID 993 1002 BY SIMILARITY.
 FT DISULFID 1005 1020 BY SIMILARITY.
 FT DISULFID 1058 1058 INTERCHAIN (PROBABLE).
 FT DISULFID 1061 1061 INTERCHAIN (PROBABLE).
 FT CARBOHYD 591 591 POTENTIAL.
 FT CARBOHYD 912 912 POTENTIAL.
 FT CARBOHYD 1398 1398 POTENTIAL.
 FT CARBOHYD 1500 1500 POTENTIAL.
 FT CARBOHYD 1571 1571 POTENTIAL.
 FT CARBOHYD 1600 1600 POTENTIAL.
 FT CARBOHYD 1737 1737 POTENTIAL.
 FT CARBOHYD 1819 1819 POTENTIAL.
 FT CARBOHYD 1886 1886 POTENTIAL.
 FT CARBOHYD 2333 2333 POTENTIAL.
 FT CARBOHYD 2509 2509 POTENTIAL.
 FT VARSPLIC 1 842 MISSING (IN ISOFORM A).
 FT VARSPLIC 843 901 MVLPRRLRLHRLGHFTETORLTGLVEGLLEASDTGSGPR
 FT AALGCLGSGEQKRVAFQRPQNHQASLWELRPS (IN
 FT ISOFORM A).
 SQ SEQUENCE 2569 AA; 282159 MW; 1AF0167A CRC32:

Query Match 89.7%; Score 26; DB 1; Length 2569;
 Best Local Similarity 83.3%; Pred. No. 2.9e-02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
 Db 19 RIVLY 24

RESULT 6
 LMA_DROME
 ID LMA_DROME STANDARD; PRT; 3712 AA.
 AC Q00174;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE LAMININ ALPHA CHAIN PRECURSOR.

GN LANA OR LAMA.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 93049203.
 RA KUSCHE-GULLBERG M., GARRISON K., MACKRELL A.J., FESSLER L.I.,
 RA FESSLER J.H.;
 RT "Laminin A chain: expression during Drosophila development and
 RT genomic sequence.";
 RL EMBO J. 11:4519-4527(1992).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE-EMBRYO;
 RX MEDLINE; 94038678.
 RA HENCHLIFFE C., GARCIA-ALONSO L., TANG J., GOODMAN C.S.;
 RT "Genetic analysis of laminin A reveals diverse functions during
 RT morphogenesis in Drosophila.";
 RN Development 118:325-337(1993).
 RN (3)
 RP SEQUENCE OF 1762-3712 FROM N.A.
 RX MEDLINE; 92078147.
 RA GARRISON K., MACKRELL A.J., FESSLER J.H.;
 RT "Drosophila laminin A chain sequence, interspecies comparison, and
 RT domain structure of a major carboxyl portion.";
 RL J. Biol. Chem. 266:22899-22904(1991).
 CC -|- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
 CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
 CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
 CC -|- FUNCTION: DIVERSE FUNCTIONS DURING MORPHOGENESIS IN DROSOPHILA.
 CC COMPLETE LOSS-OF-FUNCTION MUTATIONS LEAD TO LATE EMBRYONIC
 CC LETHALITY. CERTAIN PARTIAL LOSS-OF-FUNCTION MUTATIONS GIVE RAISE
 CC TO ESCAPER ADULTS, WHICH HAVE ROUGH EYES ASSOCIATED WITH CHANGES
 CC IN CELL FATE AND PATTERN, MISSHAPEN LEGS AND DEFECTS IN WING
 CC STRUCTURE.
 CC -|- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
 CC -|- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
 CC MEMBRANES (MAJOR COMPONENT).
 CC -|- TISSUE SPECIFICITY: NEWLY FORMED MESODERM AND LATER PROMINENTLY
 CC EXPRESSED IN HEMOCYTES, WHICH ALSO SYNTHESIZE COLLAGEN IV.
 CC -|- DEVELOPMENTAL STAGE: DURING MORPHOGENESIS, MOSTLY IN EMBRYO
 CC DEVELOPMENT AT 10-12 HOURS.
 CC -|- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -|- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
 CC -|- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -|- SIMILARITY: CONTAINS 21.5 LAMININ EGF-LIKE DOMAINS.
 CC -|- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV (DOMAIN IV) IS NOT
 CC SIMILAR TO LAMININ DOMAIN IV).
 CC -|- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M96388; AAA28662.1;
 CC EMBL; L07288; AAC37178.1;
 CC EMBL; M75882; AAA28661.1;
 CC HSSP; P02468; 1TLE.
 CC FLYBASE; FBgn0002526; LANA.
 CC PROSITE; PS00022; EGF_1; 17.
 CC PROSITE; PS01186; EGF_2; 5.
 CC PROSITE; PS01248; LAMININ_TYPE_EGF; 19.

DR PFAM; PF00052; laminin_B; 1.
DR PFAM; PF00053; laminin_EGF; 20.
DR PFAM; PF00054; laminin_G; 5.
DR PFAM; PF00055; laminin_Nterm; 1.
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 3712 LAMININ ALPHA CHAIN.
FT DOMAIN 25 272 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 273 815 10.5 X LAMININ EGF-LIKE REPEATS (DOMAIN V).
FT DOMAIN 273 332 LAMININ EGF-LIKE 1.
FT DOMAIN 333 402 LAMININ EGF-LIKE 2.
FT DOMAIN 403 447 LAMININ EGF-LIKE 3.
FT DOMAIN 448 494 LAMININ EGF-LIKE 4.
FT DOMAIN 495 540 LAMININ EGF-LIKE 5.
FT DOMAIN 541 586 LAMININ EGF-LIKE 6.
FT DOMAIN 587 631 LAMININ EGF-LIKE 7.
FT DOMAIN 632 676 LAMININ EGF-LIKE 8.
FT DOMAIN 677 731 LAMININ EGF-LIKE 9.
FT DOMAIN 732 784 LAMININ EGF-LIKE 10.
FT DOMAIN 785 815 LAMININ EGF-LIKE 11 (INCOMPLETE).
FT DOMAIN 816 1374 LAMININ EGF-LIKE 11 (INCOMPLETE).
FT DOMAIN 1375 1574 4.5 X LAMININ EGF-LIKE REPEATS (DOMAIN III B).
FT DOMAIN 1375 1420 LAMININ EGF-LIKE 12.
FT DOMAIN 1421 1465 LAMININ EGF-LIKE 13.
FT DOMAIN 1466 1513 LAMININ EGF-LIKE 14.
FT DOMAIN 1514 1564 LAMININ EGF-LIKE 15.
FT DOMAIN 1565 1574 LAMININ EGF-LIKE 16 (N-TERMINAL).
FT DOMAIN 1575 1775 LAMININ DOMAIN IV (DOMAIN IV).
FT DOMAIN 1776 2111 6.5 X LAMININ EGF-LIKE REPEATS (DOMAIN III A).
FT DOMAIN 1776 1808 LAMININ EGF-LIKE 16 (C-TERMINAL).
FT DOMAIN 1809 1858 LAMININ EGF-LIKE 17.
FT DOMAIN 1859 1916 LAMININ EGF-LIKE 18.
FT DOMAIN 1917 1969 LAMININ EGF-LIKE 19.
FT DOMAIN 1970 2016 LAMININ EGF-LIKE 20.
FT DOMAIN 2017 2063 LAMININ EGF-LIKE 21.
FT DOMAIN 2064 2111 LAMININ EGF-LIKE 22.
FT DOMAIN 2112 2697 LAMININ II AND I.
FT DOMAIN 2698 3712 5 X LAMININ G-LIKE REPEATS (DOMAIN G).
FT DOMAIN 2698 2862 LAMININ G-LIKE 1.
FT DOMAIN 2863 3048 LAMININ G-LIKE 2.
FT DOMAIN 3049 3223 LAMININ G-LIKE 3.
FT DOMAIN 3270 3296 POLY-THR.
FT DOMAIN 3334 3528 LAMININ G-LIKE 4.
FT DOMAIN 3529 3712 LAMININ G-LIKE 5.
FT DOMAIN 2178 2249 COILED COIL (POTENTIAL).
FT DOMAIN 2301 2321 COILED COIL (POTENTIAL).
FT DOMAIN 2375 2450 COILED COIL (POTENTIAL).
FT DOMAIN 2541 2676 COILED COIL (POTENTIAL).
FT DISULFID 273 282 BY SIMILARITY.
FT DISULFID 275 296 BY SIMILARITY.
FT DISULFID 298 307 BY SIMILARITY.
FT DISULFID 310 330 BY SIMILARITY.
FT DISULFID 333 342 BY SIMILARITY.
FT DISULFID 335 367 BY SIMILARITY.
FT DISULFID 370 379 BY SIMILARITY.
FT DISULFID 382 400 BY SIMILARITY.
FT DISULFID 403 414 BY SIMILARITY.
FT DISULFID 405 421 BY SIMILARITY.
FT DISULFID 423 432 BY SIMILARITY.
FT DISULFID 435 445 BY SIMILARITY.
FT DISULFID 448 460 BY SIMILARITY.
FT DISULFID 450 468 BY SIMILARITY.
FT DISULFID 470 479 BY SIMILARITY.
FT DISULFID 482 492 BY SIMILARITY.
FT DISULFID 495 507 BY SIMILARITY.
FT DISULFID 497 514 BY SIMILARITY.
FT DISULFID 516 525 BY SIMILARITY.
FT DISULFID 528 538 BY SIMILARITY.
FT DISULFID 541 553 BY SIMILARITY.
FT DISULFID 560 560 BY SIMILARITY.
FT DISULFID 562 562 BY SIMILARITY.
FT DISULFID 574 584 BY SIMILARITY.
FT DISULFID 587 599 BY SIMILARITY.
FT DISULFID 607 616 BY SIMILARITY.
FT DISULFID 619 629 BY SIMILARITY.
FT DISULFID 632 644 BY SIMILARITY.
FT DISULFID 634 650 BY SIMILARITY.
FT DISULFID 652 661 BY SIMILARITY.
FT DISULFID 664 674 BY SIMILARITY.
FT DISULFID 677 691 BY SIMILARITY.
FT DISULFID 692 700 BY SIMILARITY.
FT DISULFID 702 711 BY SIMILARITY.
FT DISULFID 714 729 BY SIMILARITY.
FT DISULFID 732 746 BY SIMILARITY.
FT DISULFID 734 753 BY SIMILARITY.
FT DISULFID 755 764 BY SIMILARITY.
FT DISULFID 767 782 BY SIMILARITY.
FT DISULFID 1375 1387 BY SIMILARITY.
FT DISULFID 1377 1394 BY SIMILARITY.
FT DISULFID 1396 1405 BY SIMILARITY.
FT DISULFID 1408 1418 BY SIMILARITY.
FT DISULFID 1421 1429 BY SIMILARITY.
FT DISULFID 1423 1436 BY SIMILARITY.
FT DISULFID 1438 1447 BY SIMILARITY.
FT DISULFID 1450 1463 BY SIMILARITY.
FT DISULFID 1466 1480 BY SIMILARITY.
FT DISULFID 1468 1487 BY SIMILARITY.
FT DISULFID 1489 1498 BY SIMILARITY.
FT DISULFID 1501 1511 BY SIMILARITY.
FT DISULFID 1514 1526 BY SIMILARITY.
FT DISULFID 1516 1533 BY SIMILARITY.
FT DISULFID 1535 1544 BY SIMILARITY.
FT DISULFID 1547 1562 BY SIMILARITY.
FT DISULFID 1859 1874 BY SIMILARITY.
FT DISULFID 1861 1885 BY SIMILARITY.
FT DISULFID 1887 1896 BY SIMILARITY.
FT DISULFID 1899 1914 BY SIMILARITY.
FT DISULFID 1917 1931 BY SIMILARITY.
FT DISULFID 1919 1938 BY SIMILARITY.
FT DISULFID 1941 1950 BY SIMILARITY.
FT DISULFID 1953 1967 BY SIMILARITY.
FT DISULFID 1970 1980 BY SIMILARITY.
FT DISULFID 1972 1987 BY SIMILARITY.
FT DISULFID 1989 1998 BY SIMILARITY.
FT DISULFID 2001 2014 BY SIMILARITY.
FT DISULFID 2017 2028 BY SIMILARITY.
Query Match 89.7%; Score 26; DB 1; Length 3712;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 RILLRY 6
Db 885 RIVLRY 890
RESULT 7
RS8_BORBU STANDARD; PRT; 132 AA.
ID RS8_BORBU
AC O51445;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 30S RIBOSOMAL PROTEIN S8.
GN RPSH OR BB0492.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RX MEDLINE; 98065943.

RA FRASER C.M., CASIENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
 RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKY E.K., GINN M.,
 RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
 RA PETERSON J., KERLAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,
 RA VAN VUOT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
 RA UTTERBACK T., WATTHEY L., McDONALD L., ARTIACH P., BOWMAN C.,
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
 RA SMITH H.O., VENTER J.C.;
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
 burgdorferi";
 RL Nature 390:580-586(1997).
 CC -1- FUNCTION: BINDS DIRECTLY TO THE CENTRAL DOMAIN OF 16S RIBOSOMAL
 CC RNA (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SRP FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE001152; AAC66850.1; -;
 DR TIGR; BB0492; -;
 DR PROSITE; PS00053; RIBOSOMAL_S8; 1.
 DR PFAM; PF00410; Ribosomal_S8; 1.
 KW Ribosomal protein.
 SQ SEQUENCE 132 AA; 14814 MW; 874CB660 CRC32;

Query Match 86.2%; Score 25; DB 1; Length 132;
 Best Local Similarity 66.7%; Pred. No. 23;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RILLRY 6
 I:|:|:|
 DB 61 RVLKY 66

RESULT 8
 VPRL_MOUSE
 ID VPRL_MOUSE STANDARD; PRT; 139 AA.
 AC P12018;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-NOV-1991 (Rel. 20, Last annotation update)
 DE IMMUNOGLOBULIN IOTA CHAIN PRECURSOR (V(PREB) PROTEIN) (FRAGMENT).
 GN VPREB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 88196069.
 RA BAUER S.R., KUDO A., MELCHERS F.;
 RT "Structure and pre-B lymphocyte restricted expression of the VpreB in
 RT humans and conservation of its structure in other mammalian
 RT species";
 RL EMBO J. 7:111-116(1988).
 CC -1- FUNCTION: ASSOCIATES WITH THE IG-MU CHAIN TO FORM A MOLECULAR
 CC COMPLEX THAT IS EXPRESSED ON THE SURFACE OF PRE-B-CELLS. THIS
 CC COMPLEX PRESUMABLY REGULATES IG GENE REARRANGEMENTS IN THE EARLY
 CC STEPS OF B-CELL DIFFERENTIATION.
 CC -1- TISSUE SPECIFICITY: ONLY EXPRESSED BY PRE-B-CELLS.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M34927; AAA61292.1; -;
 DR PIR; S00258; S00258.
 DR MIM; 146770; -;
 DR PFAM; PF00047; Ig; 1.
 KW Immunoglobulin domain; B-cell; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 139
 FT DOMAIN 20 41
 FT DOMAIN 42 56
 FT DOMAIN 57 70
 FT DOMAIN 71 81
 FT DOMAIN 82 115
 FT DISULFID 41 115
 FT NON_TER 139 139
 SQ SEQUENCE 139 AA; 15948 MW; E1DA1049 CRC32;

Query Match 86.2%; Score 25; DB 1; Length 139;
 Best Local Similarity 83.3%; Pred. No. 24;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RILLRY 6
 I:|:|:|
 DB 66 RFLRY 71

RESULT 9
 VPRL_MOUSE
 ID VPRL_MOUSE STANDARD; PRT; 142 AA.
 AC P13372;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IMMUNOGLOBULIN IOTA CHAIN PRECURSOR (V(PREB)1 PROTEIN).
 GN VPREB1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6 X DBA/2J;
 RX MEDLINE; 88029315.
 RA KUDO A., MELCHERS F.;
 RT "A second gene, VpreB in the lambda 5 locus of the mouse, which
 RT appears to be selectively expressed in pre-B lymphocytes";
 RL EMBO J. 6:2267-2272(1987).
 CC -1- FUNCTION: ASSOCIATES WITH THE IG-MU CHAIN TO FORM A MOLECULAR
 CC COMPLEX THAT IS EXPRESSED ON THE SURFACE OF PRE-B-CELLS. THIS
 CC COMPLEX PRESUMABLY REGULATES IG GENE REARRANGEMENTS IN THE EARLY
 CC STEPS OF B-CELL DIFFERENTIATION.
 CC -1- TISSUE SPECIFICITY: ONLY EXPRESSED BY PRE-B-CELLS.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X05556; CAA29071.1; -;
 DR EMBL; X05557; CAA29072.1; -;
 DR PIR; A28344; A28344.
 DR HSP; P01607; IREI.
 DR MGD; MGI:98936; VPREB1.
 DR PFAM; PF00047; Ig; 1.
 KW Immunoglobulin domain; B-cell; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 142
 IMMUNOGLOBULIN IOTA CHAIN.

FT DOMAIN 20 41 FRAMEWORK 1.
FT DOMAIN 42 56 COMPLEMENTARITY-DETERMINING 1.
FT DOMAIN 57 70 FRAMEWORK 2.
FT DOMAIN 71 81 COMPLEMENTARITY-DETERMINING 2.
FT DOMAIN 82 115 FRAMEWORK 3.
FT DISULFID 41 115 BY SIMILARITY.
SQ SEQUENCE 142 AA; 16125 MW; 45C58A1B CRC32;

Query Match 86.2%; Score 25; DB 1; Length 142;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RILLY 6
Db 66 RFLRY 71

RESULT 10
VPR2_MOUSE STANDARD; PRT; 142 AA.
AC P13373;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IMMUNOGLOBULIN OMEGA CHAIN PRECURSOR (V(PREB)2 PROTEIN).
GN VPREB2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6 X DBA/2J;
RX MEDLINE; 88029315.
RA KUDO A., MELCHERS F.;
RT "A second gene, Vpreb in the lambda 5 locus of the mouse, which appears to be selectively expressed in pre-B lymphocytes.";
RL EMBO J. 6:2267-2272(1987).
CC -1- FUNCTION: ASSOCIATES WITH THE IG-MU CHAIN TO FORM A MOLECULAR COMPLEX PRESUMABLY REGULATES IG GENE REARRANGEMENTS IN THE EARLY STEPS OF B-CELL DIFFERENTIATION.
CC -1- TISSUE SPECIFICITY: ONLY EXPRESSED BY PRE-B-CELLS.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
CC EMBL; X05563; CAA29077.1; -
CC DR PIR; B28344; B28344.
CC DR HSSP; P01607; 1REI.
CC DR MGD; MGI:98937; VPREB2.
CC DR PFAM; PF00047; Ig; 1.
CC Immunoglobulin domain; B-cell; Signal.
CC SIGNAL 1 19 POTENTIAL.
CC FT CHAIN 20 142 IMMUNOGLOBULIN OMEGA CHAIN.
CC FT DOMAIN 20 41 FRAMEWORK 1.
CC FT DOMAIN 42 56 COMPLEMENTARITY-DETERMINING 1.
CC FT DOMAIN 57 70 FRAMEWORK 2.
CC FT DOMAIN 71 81 COMPLEMENTARITY-DETERMINING 2.
CC FT DOMAIN 82 115 FRAMEWORK 3.
CC FT DISULFID 41 115 BY SIMILARITY.
SQ SEQUENCE 142 AA; 16052 MW; 6AE40A3B CRC32;

Query Match 86.2%; Score 25; DB 1; Length 142;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RILLY 6
Db 66 RFLRY 71

RESULT 11
MEFB_HUMAN STANDARD; PRT; 365 AA.
AC Q02080;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MYOCYTE-SPECIFIC ENHANCER FACTOR 2B (SERUM RESPONSE FACTOR-LIKE PROTEIN 2) (XMEF2) (RSRFR2).
GN MEF2B OR XMEF2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-HEART, AND SKELETAL MUSCLE;
RX MEDLINE; 92387551.
RA YU Y.-T., BREITBART R.E., SMOOT L.B., LEE Y., MAHDAVI V.,
RA NADAL-GINARD B.;
RT "Human myocyte-specific enhancer factor 2 comprises a group of tissue-restricted MADS box transcription factors.";
RL Genes Dev. 6:1783-1798(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA;
RX MEDLINE; 92084105.
RA POLLOCK R., TREISMAN R.;
RT "Human SRF-related proteins: DNA-binding properties and potential regulatory targets.";
RL Genes Dev. 5:2327-2341(1991).
RN [3]
RP SEQUENCE FROM N.A.
RA LAMERDIN J.E.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: TRANSCRIPTION FACTOR WHICH BINDS SPECIFICALLY TO THE MEF2 ELEMENT PRESENT IN THE REGULATORY REGIONS OF MANY MUSCLE-SPECIFIC GENES. ACTIVATES TRANSCRIPTION VIA THIS ELEMENT. MAY BE INVOLVED IN MUSCLE-SPECIFIC AND/OR GROWTH FACTOR-RELATED TRANSCRIPTION.
CC -1- SUBUNIT: HETERODIMER.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SKELETAL AND CARDIAC MUSCLE AND BRAIN.
CC -1- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS. MEF2 SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
CC EMBL; X68502; CAA48515.1; -
CC DR EMBL; X63380; CAA44978.1; -
CC DR EMBL; AC002126; AAB86982.1; -
CC DR HSSP; P11831; 1SR5.
CC DR MIM; 600661; -
CC DR PROSITE; PS00350; MADS_BOX_1; 1.
CC DR PROSITE; PS50066; MADS_BOX_2; 1.
CC DR PFAM; PF00319; SRF-TF; 1.
CC Transcription regulation; Nuclear protein; DNA-binding; Activator;
CC Multigene family.
CC FT DOMAIN 3 57 MADS.
CC FT DNA_BIND 58 86 MEF2-TYPE (POTENTIAL).
CC FT DOMAIN 4 31 LYS-RICH (BASIC).
CC FT DOMAIN 5 31

SQ SEQUENCE 365 AA; 38638 MW; 59008C81 CRC32;

Query Match 86.2%; Score 25; DB 1; Length 365;

Best Local Similarity 66.7%; Pred. No. 67;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6

Db 64 RVLKY 69

RESULT 12

PRS2_METTM

ID PRS2_METTM STANDARD; PRT; 372 AA.

AC P42811;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE PUTATIVE 26S PROTEASE REGULATORY SUBUNIT HOMOLOG MTH1011.

OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).

OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;

CC Methanobacterium.

RN [1]

RP SEQUENCE FROM N.A.

RA HUNGERER C., WEISS D., THAUER R.K., JAHN D.;

RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: THE 26S PROTEASE IS INVOLVED IN THE ATP-DEPENDENT

DEGRADATION OF UBQUITINATED PROTEINS. THE REGULATORY (OR ATPASE)

CC COMPLEX CONFERS ATP DEPENDENCY AND SUBSTRATE SPECIFICITY TO THE

CC 26S COMPLEX (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X83691; CAA58665.1; -

DR PROSITE; PS00674; AAA; FALSE_NEG.

DR PFAM; PF00004; AAA; 1.

KW Proteasome; ATP-binding.

FT NP_BIND 164 171 ATP (POTENTIAL).

SQ SEQUENCE 372 AA; 42689 MW; 90B2F3CB CRC32;

Query Match

Best Local Similarity 66.7%; Pred. No. 68;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6

Db 140 RIIMY 145

RESULT 13

TRSA_STRAM

ID TRSA_STRAM STANDARD; PRT; 415 AA.

AC Q07197; Q07199;

DT 01-NOV-1997 (Rel. 35, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE TRANSFER PROTEIN TRSA.

GN TRSA.

OS Streptomyces ambifaciens.

OG Plasmid pSAM2.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 23877;
RX MEDLINE; 93374848.
RA HAGEGE J., PERNODET J.L., SEZONOV G., GERBAUD C., FRIEDMANN A.,
RA GUERINEAU M.;
RT "Transfer functions of the conjugative integrating element pSAM2 from
RT Streptomyces ambifaciens: characterization of a kil-kor system
RT associated with transfer";
RL J. Bacteriol. 175:5529-5538(1993).
RN [2]

RP REVISIONS.

RA SEZONOV G.;

RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE FTSK/SPOIIIE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; Z19593; CAA79645.1; -

DR PFAM; PF01580; FtsK_SpoIIIE; 1.

KW Plasmid; ATP-binding.

FT NP_BIND 145 152 ATP (POTENTIAL).

SQ SEQUENCE 415 AA; 45263 MW; 468A7AA2 CRC32;

Query Match

Best Local Similarity 86.2%; Score 25; DB 1; Length 415;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6

Db 2 RVLRY 7

RESULT 14

FEAB_ECOLI

ID FEAB_ECOLI STANDARD; PRT; 499 AA.

AC P80668; P46884; P77637; O32557;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE PHENYLACETALDEHYDE DEHYDROGENASE (EC 1.2.1.39) (PAD).

GN FEAB OR PADA OR MAOB.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-9.

RC STRAIN-K12;

RX MEDLINE; 97195795.

RA HANLON S.P., HILL T.K., FLAVELL M.A., STRINGFELLOW J.M., COOPER R.A.;

RT "2-phenylethylamine catabolism by Escherichia coli K-12: gene

RT organization and expression";

RL Microbiology 143:513-518(1997).

RN [2]

RP REVISIONS.

RA STRINGFELLOW J.M.;

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.

RC STRAIN-W / ATCC 11105;

RX MEDLINE; 97263463.

RA FERRANDEZ A., PRIETO M.A., GARCIA J.L., DIAZ E.;

RT "Molecular characterization of PadA, a phenylacetaldehyde

RT dehydrogenase from Escherichia coli.;

RL FEBS Lett. 406:23-27(1997).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;

```

RX MEDLINE: 97426617.
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GORDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RA AIGA H., BABA T., FUJITA K., HAYASHI K., HONJO A., HORIUCHI T.,
RA IKEMOTO K., INADA T., ISONO K., ISONO S., ITOH T., KANAI K., KASAI H.,
RA KASHIMOTO K., KIM S., KIMURA S., KITAGAWA M., KITAKAWA M., MAKINO K.,
RA MASUDA S., MIKI T., MIYABUCHI K., MORI H., MOTOMURA K., NAKAMURA Y.,
RA NASHIMOTO H., NISHIO Y., OSHINA T., SAITO N., SAMPEI G., SEKI Y.,
RA TAGAMI H., TAKEMOTO K., WADA C., YAMAMOTO Y., YANO M.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [6]
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
RC STRAIN-K12 / W3110;
RA AZAKAMI H., YAMASHITA M., ROH J.-H., SUZUKI H., KUMAGAI H.,
RA MUROOKA Y.;
RT "Nucleotide sequence of the gene for monoamine oxidase (maoa) from
RT Escherichia coli.";
RL J. Ferment. Bioeng. 77:315-319(1994).
CC -!- FUNCTION: ACTS ALMOST EQUALLY WELL ON PHENYLACETALDEHYDE, 4-
CC HYDROXYPHENYLACETALDEHYDE AND 3,4-DIHYDROXYPHENYLACETALDEHYDE.
CC -!- CATALYTIC ACTIVITY: PHENYLACETALDEHYDE + NAD(+) + H(2)O =
CC PHENYLACETATE + NADH.
CC -!- COFACTOR: PREFERENCES OVER NADP.
CC -!- PATHWAY: INITIAL STEPS OF 2-PHENYLETHYLAMINE CATABOLISM.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X99402; CAA67780.1; -.
CC EMBL; X97453; CAA66106.1; -.
CC EMBL; AE000235; AAC74467.1; ALT_INIT.
CC EMBL; D90776; CAB20903.1; ALT_INIT.
CC EMBL; D90777; CAB20911.1; ALT_INIT.
CC EMBL; D23670; -; NOT_ANNOTATED_CDS.
CC HSP; P20000; 1A4Z.
CC ECOGENE; EG13179; FEAB.
CC PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
CC PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
CC PFAM; PF00171; aldedh; 1.
CC Oxidoreductase; NAD.
CC NP_BIND 250 255 NAD (ADP PART) (BY SIMILARITY).
CC ACT_SITE 272 272 BY SIMILARITY.
CC FT ACT_SITE 306 306 BY SIMILARITY.
CC FT CONFLICT 265 265 H -> R (IN REF. 3).
CC FT CONFLICT 484 499 DWLGGCETKSCVCVRY -> PLGRAGC (IN REF. 1).
CC SEQUENCE 499 AA; 53699 MW; FFC195E0 CRC32;

Query Match 86.2%; Score 25; DB 1; Length 499;
Best Local Similarity 83.3%; Pred. No. 93;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RILEY 6
| | | | |
Db 85 RILLRF 90

RESULT 15
MEF2_DROME

```

```

ID AC MEF2_DROME STANDARD; PRT; 515-AA.
AD P40791;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MYOCYTE-SPECIFIC ENHANCER FACTOR 2.
GN MEF2.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-EMBRYO;
RX MEDLINE; 94261646.
RA LILLY B., GALEWSKY S., FIRULLI A.B., SCHULZ R.A., OLSON E.N.;
RT "D-MEF2: a MADS box transcription factor expressed in differentiating
RT mesoderm and muscle cell lineages during Drosophila embryogenesis.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:5662-5666(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94329547.
RA NGUYEN H.T., BODMER R., ABMAYR S.M., MCDERMOTT J.C., SPOEREL N.A.;
RT "D-mef2: a Drosophila mesoderm-specific MADS box-containing gene with
RT a biphasic expression profile during embryogenesis.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:7520-7524(1994).
CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT COULD BE A KEY PLAYER IN EARLY
CC MESODERM DIFFERENTIATION AND MAY BE REQUIRED FOR SUBSEQUENT CELL
CC FATE SPECIFICATIONS WITHIN THE SOMATIC AND VISCERAL/HEART
CC MESODERMAL LAYERS.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -!- TISSUE SPECIFICITY: MESODERM.
CC -!- DEVELOPMENTAL STAGE: FIRST DETECTABLE IN THE PRESUMPTIVE MESODERM
CC AT LATE CELLULAR BLASTODERM STAGE. EXPRESSED IN ALL PRESUMPTIVE
CC MESODERM PRIOR TO THE SPLITTING PROCESS THAT GENERATES THE SOMATIC
CC AND VISCERAL/ HEART MESODERM. AFTER THE SUBDIVISION, IT IS FOUND
CC IN BOTH THE SOMATIC AND THE VISCERAL/HEART MESODERM.
CC -!- INDUCTION: TWI ACTIVITY IS REQUIRED FOR MEF2 EXPRESSION. SNA
CC ACTIVITY IS NEEDED FOR MAINTAINING MEF2 EXPRESSION.
CC -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC FACTORS. MEF2 SUBFAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U03292; AAA19957.1; -.
CC EMBL; U07422; AAA20463.1; -.
CC HSP; P11833; ISRS.
CC FLYBASE; FBgn0011656; Mef2.
CC PROSITE; PS00350; MADS_BOX_1; 1.
CC PROSITE; PS50066; MADS_BOX_2; 1.
CC PFAM; PF00319; SRF-TF; 1.
CC Transcription regulation; DNA-binding; Nuclear protein;
CC Developmental protein; Differentiation; Activator.
CC DOMAIN 3 57 MADS.
CC FT DNA_BIND 58 86 MEF2-TYPE (POTENTIAL).
CC FT DOMAIN 380 386 POLY-SER.
CC FT DOMAIN 390 396 POLY-GLY.
CC FT DOMAIN 421 428 POLY-GLY.
CC FT DOMAIN 484 495 POLY-GLN.
CC FT CONFLICT 338 338 S -> R (IN REF. 2).
CC FT CONFLICT 364 377 PAVISRIAMVPRAG -> ASGHQNSNGSTGS (IN
CC REF. 2).
CC SEQUENCE 515 AA; 54289 MW; 0D46F188 CRC32;

Query Match 86.2%; Score 25; DB 1; Length 515;
Best Local Similarity 66.7%; Pred. No. 96;

```

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLLRY 6
|:|:|
Db 64 RLLRY 69

Search completed: February 8, 2000, 00:59:46
Job time: 3775 sec

| Result No. | Query | | | DB | ID | Description |
|------------|-------|-------|--------|----|--------|---------------------|
| | Score | Match | Length | | | |
| 1 | 29 | 100.0 | 133 | 2 | Q9Z903 | Q9Z903 chlamydia p |
| 2 | 29 | 100.0 | 458 | 2 | Q84822 | Q84822 chlamydia t |
| 3 | 29 | 100.0 | 458 | 2 | Q92601 | Q92601 chlamydia t |
| 4 | 29 | 100.0 | 532 | 2 | P72007 | P72007 mycobacteri |
| 5 | 28 | 96.6 | 616 | 1 | Q51714 | Q51714 myrococcus |
| 6 | 27 | 93.1 | 323 | 5 | O16335 | O16335 caenorhabdi |
| 7 | 27 | 93.1 | 332 | 2 | Q84491 | Q84491 chlamydia t |
| 8 | 27 | 93.1 | 600 | 2 | Q9ZG01 | Q9ZG01 yersinia pe |
| 9 | 27 | 93.1 | 600 | 2 | Q9X915 | Q9X915 yersinia en |
| 10 | 27 | 93.1 | 777 | 4 | O43162 | O43162 homo sapien |
| 11 | 26 | 89.7 | 42 | 11 | O54741 | O54741 mus musculus |
| 12 | 26 | 89.7 | 333 | 10 | Q9ZRS3 | Q9ZRS3 medicago sa |
| 13 | 26 | 89.7 | 383 | 5 | Q18310 | Q18310 caenorhabdi |
| 14 | 26 | 89.7 | 502 | 2 | O53551 | O53551 mycobacteri |
| 15 | 26 | 89.7 | 540 | 2 | O87689 | O87689 bacillus me |
| 16 | 26 | 89.7 | 545 | 11 | Q9WV73 | Q9WV73 mus musculus |
| 17 | 26 | 89.7 | 617 | 1 | Q9YD14 | Q9YD14 aeropyrum p |
| 18 | 26 | 89.7 | 768 | 2 | Q86766 | Q86766 streptomyce |
| 19 | 26 | 89.7 | 1486 | 4 | O14637 | O14637 homo sapien |
| 20 | 25 | 86.2 | 100 | 6 | O77624 | O77624 bos taurus |

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-D/UW-3/CX;
 RA STEPHENS R.S., KALMAN S., LAMMEL C.J., FAN J., MARATHE R., ARAVIND L.,
 RA MITCHELL W.P., OLINGER L., TATISOV R.L., ZHAO Q., KOONIN E.V.,
 RA DAVIS R.W.;
 RT "Genome Sequence of an Obligate Intracellular Pathogen of Humans:
 RT Chlamydia trachomatis";
 RN Science 0:0-0(1998).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-D/UW-3/CX;
 RA STEPHENS R.S., KALMAN S., LAMMEL C.J., FAN J., MARATHE R., ARAVIND L.,
 RA MITCHELL W.P., OLINGER L., TATISOV R.L., ZHAO Q., KOONIN E.V.,
 RA DAVIS R.W.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE001354; AAC68412.1; -;
 DR PROSITE: PS00710; PGM_PMM; 1.
 DR PFAM: PF00408; PGM_PMM; 1.
 SQ SEQUENCE 458 AA; 49391 MW; 8513D3D6 CRC32;

Query Match 100.0%; Score 29; DB 2; Length 458;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RILLY 6
 Db 413 RILLY 418

RESULT 3
 Q92601 PRELIMINARY; PRT; 458 AA.
 AC Q92601;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE PHOSPHOGLUCOMUTASE.
 GN PGM.
 OS Chlamydia pneumoniae.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia phila.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CWL029;
 RA KALMAN S., MITCHELL W., MARATHE R., LAMMEL C., FAN J., OLINGER L.,
 RA GRIMWOOD J., DAVIS R.W., STEPHENS R.S.;
 RT "Comparative Genomes of Chlamydia pneumoniae and C. trachomatis";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE001677; AAD19103.1; -;
 DR PROSITE: PS00710; PGM_PMM; 1.
 SQ SEQUENCE 458 AA; 49852 MW; 68C5A8CB CRC32;

Query Match 100.0%; Score 29; DB 2; Length 458;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RILLY 6
 Db 413 RILLY 418

RESULT 4
 P72007 PRELIMINARY; PRT; 532 AA.
 ID P72007; O08153;
 AC P72007;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE HYPOTHETICAL 57.4 KD PROTEIN.
 GN MTCY28.13C.
 OS Mycobacterium tuberculosis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RA OLIVER K., MCLEAN J., HARRIS D.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z95890; CAB09316.1; -;
 DR PROSITE: PS00455; AMP-BINDING; 1.
 DR PFAM: PF00501; AMP-binding; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 532 AA; 57413 MW; 7FC653FB CRC32;

Query Match 100.0%; Score 29; DB 2; Length 532;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RILLY 6
 Db 419 RILLY 424

RESULT 5
 Q51714 PRELIMINARY; PRT; 616 AA.
 AC Q51714;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE PROLYL ENDOPEPTIDASE.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DSM 3638;
 RX MEDLINE; 95129900.
 RA ROBINSON K.A., BARTLEY D.A., ROBB F.T., SCHREIER H.J.;
 RT "A gene from the hyperthermophile Pyrococcus furiosus whose deduced
 RT product is homologous to members of the prolyl oligopeptidase family
 RT of proteases";
 RL Gene 152:103-106(1995).
 DR EMBL: U08343; AAA73423.1; -;
 DR PFAM: PF00326; Peptidase_S9; 1.
 DR PRINTS: PR00862; PROLIGOPTASE.
 SQ SEQUENCE 616 AA; 70867 MW; 86AC623A CRC32;

Query Match 96.6%; Score 28; DB 1; Length 616;
 Best Local Similarity 83.3%; Pred. No. 79;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RILLY 6
 Db 334 RILLY 339

RESULT 6
 O16335 PRELIMINARY; PRT; 323 AA.
 ID O16335;
 AC O16335;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
 DE C09H5.8 PROTEIN.
 GN C09H5.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;

OC Rhabdittina; Rhabditoidea; Rhabdittidae; Peloderinae; Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE; 94150718.
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., JIER M., JOHNSTON L.,
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTECK L., WILKINSON-SPROAT J., WOLDMAN P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans";
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA LE T. T., WATERSTON R.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF016433; AAB5392.1; -;
 DR PFAM; PF01461; 7tm4; 1.
 SQ SEQUENCE 323 AA; 37314 MW; 0AC10232 CRC32;

Query Match 93.1%; Score 27; DB 5; Length 323;
 Best Local Similarity 83.3%; Pred. No. 73;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RILLY 6
 Db 177 RLLRY 182

RESULT 7
 O84491 ID O84491 PRELIMINARY; PRT; 332 AA.
 AC O84491;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE HYPOTHETICAL 38.6 KD PROTEIN.
 GN CT484.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-D/UW-3/CX;
 RA STEPHENS R.S., KALMAN S., LAMMEL C.J., FAN J., MARATHE R., ARAVIND L.,
 RA MITCHELL W.P., OLINGER L., TATISOV R.L., ZHAO Q., KOONIN E.V.,
 RA DAVIS R.W.;
 RT "Genome Sequence of an Obligate Intracellular Pathogen of Humans:
 Chlamydia trachomatis";
 RL Science 0:0-0(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-D/UW-3/CX;
 RA STEPHENS R.S., KALMAN S., LAMMEL C.J., FAN J., MARATHE R., ARAVIND L.,
 RA MITCHELL W.P., OLINGER L., TATISOV R.L., ZHAO Q., KOONIN E.V.,
 RA DAVIS R.W.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE001322; AAC68084.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 332 AA; 38649 MW; B5C59CE9 CRC32;

Query Match 93.1%; Score 27; DB 2; Length 332;
 Best Local Similarity 83.3%; Pred. No. 75;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RILLY 6
 Db 19 RLLRY 24

RESULT 8
 O9ZG01 ID O9ZG01 PRELIMINARY; PRT; 600 AA.
 AC O9ZG01;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE YBTP.
 GN YBTP.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KIM6+;
 RX MEDLINE; 99035519.
 RA GEHRING A.M., DEMOLL E., FETHERSTON J.D., MORI I., MAYHEW G.F.,
 RA BLATTNER F.R., WALSH C.T., PERRY R.D.;
 RT "Iron acquisition in plague: modular logic in enzymatic biogenesis of
 yersiniabactin by Yersinia pestis";
 RL Chem. Biol. 5:573-586(1998).
 DR EMBL; AF091251; AAC69585.1; -;
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Transport.
 SQ SEQUENCE 600 AA; 66284 MW; D8692F0B CRC32;

Query Match 93.1%; Score 27; DB 2; Length 600;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RILLY 6
 Db 388 RLLRY 393

RESULT 9
 O9X915 ID O9X915 PRELIMINARY; PRT; 600 AA.
 AC O9X915;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE LIPOPROTEIN INNER MEMBRANE ABC-TRANSPORTER, IRP6.
 GN IRP6.
 OS Yersinia enterocolitica.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA RAKIN A.V., PELLUDAT C., NOELTING C., SCHUBERT S., JACOBI C.,
 RA HEESEMAN J.;
 RT "Common and specific characteristics of the high-pathogenicity island
 of Yersinia enterocolitica";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ322668; CAB46573.1; -;
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KW Lipoprotein; ATP-binding; Transport.
 SQ SEQUENCE 600 AA; 66413 MW; E074027F CRC32;

Query Match 93.1%; Score 27; DB 2; Length 600;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RILLY 6
 Db 388 RLLRY 393

```

RESULT 10
ID O43162 PRELIMINARY; PRT; 777 AA.
AC O43162;
DT 01-JUN-1998 (T-EMBLrel. 06, Created)
DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)
DT 01-AUG-1998 (T-EMBLrel. 07, Last annotation update)
DE KIAA0435.
GN KIAA0435.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA ISHIKAWA K., NAGASE T., NAKAJIMA D., SEKI N., OHIRA M., MIYAJIMA N.,
RA TANAKA A., KOTANI H., NOMURA N., OHARA O.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB007895; BAA23708.1; -.
DR HSP: 777 AA; 86654 MW; 998863D5 CRC32;
SQ SEQUENCE 777 AA; 86654 MW; 998863D5 CRC32;

Query Match 93.1%; Score 27; DB 4; Length 777;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
DB 179 RILLY 184

RESULT 11
ID O54741 PRELIMINARY; PRT; 42 AA.
AC O54741;
DT 01-JUN-1998 (T-EMBLrel. 06, Created)
DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)
DT 01-NOV-1999 (T-EMBLrel. 12, Last annotation update)
DE LAMININ, ALPHA 3 (LAMININ 5 ALPHA3C CHAIN) (FRAGMENT).
GN LAMA3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SV129;
RX MEDLINE: 97400527.
RA FERRIGNO O., VIROLLE T., GALLIANO M.F., CHAUVIN N., ORTONNE J.P.,
RA MENEGUZZI G., ABERDAM D.;
RT "Murine laminin alpha3a and alpha3b isoform chains are generated by
RT usage of two promoters and alternative splicing.";
RL J. Biol. Chem. 272:20502-20508(1997).
DR EMBL: Y08850; CAA70073.1; -.
DR MGD: MG1:99909; Lama3.
DR NON_TER: 42
FT SEQUENCE 42 AA; 4872 MW; 42FE6150 CRC32;

Query Match 89.7%; Score 26; DB 11; Length 42;
Best Local Similarity 83.3%; Pred. No. 17;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
DB 26 RILLY 31

RESULT 12
Q92R53 PRELIMINARY; PRT; 333 AA.
ID Q92R53
AC Q92R53;

```

```

DT 01-MAY-1999 (T-EMBLrel. 10, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
DT 01-NOV-1999 (T-EMBLrel. 12, Last annotation update)
DE ANNEXIN-LIKE PROTEIN.
GN NANN.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
OC Medicago.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RA3;
RX MEDLINE: 98388649.
RA KOVACS I., AYAYDIN F., OBERSCHALL A., IPACS I., BOTTKA S., PONGOR S.,
RA DUDITS D., CSORDAS TOTH E.;
RT "Immunolocalization of a novel annexin-like protein encoded by a
RT stress and abscisic acid responsive gene in alfalfa.";
RL Plant J. 15:185-197(1998).
DR EMBL: Y11348; CAA72183.1; -.
DR HSP: P13214; LAOW.
DR SEQUENCE 333 AA; 37934 MW; 62598772 CRC32;

Query Match 89.7%; Score 26; DB 10; Length 333;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6
DB 146 KILLY 151

RESULT 13
ID Q18310 PRELIMINARY; PRT; 383 AA.
AC Q18310;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (T-EMBLrel. 08, Last annotation update)
DE CODED FOR BY C. ELEGANS CDNA CM2082.
GN C29F5.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentes; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE: 94150718.
RA WILSON R., AINSKOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRATON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPOAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA MILLER N.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U23524; AAC46820.1; -.

```

SQ SEQUENCE 383 AA; 43263 MW; EC82938A CRC32;

Query Match 89.7%; Score 26; DB 5; Length 383;

Best Local Similarity 83.3%; Pred. NO. 1.5e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6

|||||

DB 200 RILRY 205

RESULT 14

OS3551

ID OS3551 PRELIMINARY; PRT; 502 AA.

AC OS3551; (TREMELrel. 06, Created)

DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)

DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)

DT 01-NOV-1999 (TREMELrel. 12, Last annotation update)

DE PUTATIVE COA-LIGASE.

GN MT023.13.

OS Mycobacterium tuberculosis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-H37RV;

RA BADOCK K., CHURCHER C.M.;

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-H37RV;

RA COLE S.T., PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-H37RV;

RX MEDLINE; 96181548.

RA PHILIPP W.J., POULET S., EIGLMEIER K., PASCOPELLA L.;

RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.;

RA COLE S.T.;

RT "An integrated map of the genome of the tubercle bacillus,

RT Mycobacterium tuberculosis H37Rv, and comparison with Mycobacterium

RT leprae.;"

RL Proc. Natl. Acad. Sci. U.S.A. 93:3132-3137(1996).

DR EMBL; AL022022; CAA17743.1;

DR HSSP; P08659; ILCI.

DR PROSITE; PS00455; AMP_BINDING; 1.

DR PFAM; PF00501; AMP-binding; 1.

KW Ligase.

SQ SEQUENCE 502 AA; 53738 MW; D3CF89C5 CRC32;

Query Match

Best Local Similarity 89.7%; Score 26; DB 2; Length 502;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6

|||||

DB 406 RILRY 411

RESULT 15

OS87689

ID OS87689 PRELIMINARY; PRT; 540 AA.

AC OS87689;

DT 01-NOV-1998 (TREMELrel. 08, Created)

DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)

DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)

DE PRECORIN-3 METHYLASE.

GN CBIH60.

OS Bacillus megaterium.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus..

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-DSM 509;

RX MEDLINE; 98416126.

RA RAUX E., LANOIS A., WARREN M.J., RAMBACH A., THERMES C.;

RT "Cobalamin (vitamin B12) biosynthesis: identification and

RT characterization of a Bacillus megaterium cobi operon.;"

RL Biochem. J. 335:159-166(1998).

DR EMBL; AJ000758; CAA04307.1;

DR PFAM; PF00590; PF_methylase; 1.

KW Methyltransferase.

SQ SEQUENCE 540 AA; 59510 MW; BA79DD3D CRC32;

Query Match

Best Local Similarity 89.7%; Score 26; DB 2; Length 540;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILLY 6

|||||

DB 182 RILLY 187

Search completed: February 8, 2000, 13:17:31

Job time: 32480 sec

THIS PAGE BLANK (USPTO)

| Sequence | Strd Orig | ZScore | Escore | Len | Documentation |
|---------------------|-----------|--------|---------|------|------------------------------------|
| 9b_sts:HUMSWX2 | + 29.00 | 127.30 | 45.56 | 130 | I L15224 Human chromosome x sts |
| 9b_pr2:PTRRGC | + 29.00 | 125.34 | 58.55 | 166 | I X85723 P.trogodytes mRNA for T |
| 9b_p12:AF049680 | - 29.00 | 123.69 | 72.33 | 204 | I AF049680 Neotyphodium lolii str |
| 9b_p12:AF049679 | - 29.00 | 123.46 | 74.51 | 210 | I AF049679 Epichloe typhina x Ned |
| 9b_p12:AF049681 | - 29.00 | 123.46 | 74.51 | 210 | I AF049681 Epichloe typhina strain |
| 9b_pr1:HSAR203 | + 29.00 | 120.61 | 107.42 | 300 | I AJ008203 Homo sapiens mRNA for |
| 9b_pr2:HSAR244960 | + 29.00 | 119.15 | 129.51 | 360 | I AJ244960 Homo sapiens mRNA for |
| 9b_sts:GLA238 | - 29.00 | 118.31 | 144.29 | 400 | I GLI238 human STS SHGC-7374 clon |
| 9b_sts:KLAJ3903 | - 29.00 | 117.62 | 157.62 | 436 | I G1229903 Kluyveromyces lactis di |
| 9b_pr3:AF062148 | - 29.00 | 117.53 | 159.48 | 441 | I AF062148 Homo sapiens clone tcd |
| 9b_pr3:AF056445 | - 29.00 | 117.49 | 160.22 | 441 | I AF056445 Homo sapiens clone tcd |
| 9b_sts:G23069 | - 29.00 | 117.28 | 164.67 | 455 | I G23069 human STS WI-12710 sequ |
| 9b_sts:KLAJ9696 | + 29.00 | 117.28 | 164.67 | 455 | I AJ229696 Kluyveromyces lactis di |
| 9b_pr1:HAUS6781 | - 29.00 | 115.23 | 214.21 | 588 | I U56781 Hampea appendiculis S.8 |
| 9b_pr2:CNS01995 | - 29.00 | 114.60 | 232.16 | 636 | I LL11345 Botrytis cinerea strain |
| 9b_sts:GS1197 | - 29.00 | 113.94 | 252.77 | 691 | I GS1197 SHGC-80654 Human Hom |
| 9b_p11:GGZ8FN | - 29.00 | 112.87 | 289.98 | 790 | I U17159 Gaemannomyces graminis |
| 9b_p11:PSU17158 | - 29.00 | 112.87 | 289.98 | 790 | I U17158 Philophora sp. 265 ribo |
| 9b_sts:G31223 | - 29.00 | 112.64 | 298.64 | 813 | I G31223 sp184k17-Sp6 Human (A.G |
| 9b_p11:TR25SRNA | - 29.00 | 112.60 | 300.15 | 817 | I Z17580 Trichechira reesei 25S r |
| 9b_p11:TRRRH | - 29.00 | 112.60 | 300.15 | 817 | I Z17580 Trichechira reesei 25S r |
| 9b_p11:SCDYL098C | - 29.00 | 111.43 | 348.85 | 946 | I Z74146 S.cerevisiae chromosomes |
| 9b_bal:EFLGS24B | - 29.00 | 109.46 | 449.02 | 1210 | I AJO00042 Enterococcus faecalis |
| 9b_inl:SCUTB12X | + 29.00 | 108.86 | 484.83 | 1304 | I LI3787 Styela clava alpha-musc |
| 9b_inl:SCUTB34X | + 29.00 | 108.72 | 493.60 | 1327 | I LI3780 Styela clava alpha-musc |
| 9b_inl:SCMUSACT | + 29.00 | 108.67 | 497.04 | 1336 | I X61040 S.clava mRNA for muscle |
| 9b_ov:XLIQVH6 | - 29.00 | 108.42 | 512.69 | 1377 | I X36860 X.laevis Vh VI gene for |
| 9b_ov:TRSH800221 | - 29.00 | 108.19 | 528.35 | 1418 | I AL050144 Homo sapientis mRNA; ch |
| 9b_bal:SPIL1826E | - 29.00 | 107.62 | 568.13 | 1522 | I Y18026 Streptococcus plutanima |
| 9b_p11:GGU17161 | - 29.00 | 107.10 | 607.59 | 1625 | I U17161 Gaemannomyces graminis |
| 9b_p11:GGU17160 | - 29.00 | 107.03 | 612.96 | 1639 | I U17160 Gaemannomyces graminis |
| 9b_p11:ATH011528 | + 29.00 | 106.08 | 692.48 | 1846 | I AT011528 Arabidopsis thaliana |
| 9b_ov:GECP521A | + 29.00 | 104.98 | 797.31 | 2118 | I M2036 Gecko gekko green sensil |
| 9b_ph:BPA118PLP | + 29.00 | 104.98 | 797.31 | 2118 | I X95008 Bacteriophage A118 hol |
| 9b_p11:ATAAT2E | - 29.00 | 104.90 | 805.42 | 2139 | I X94626 A.thaliana mRNA for AA |
| 9b_bal:ECRFBG | - 29.00 | 104.63 | 833.63 | 2212 | I X59852 E.coli rfb gene, 6/1993 |
| 9b_bal:BACASPKN | - 29.00 | 104.59 | 837.88 | 2223 | I M93419 Bacillus sp. aspartokin |
| 9b_p11:ATHATWEKK1 | + 29.00 | 104.03 | 900.95 | 2386 | I D50468 Arabidopsis thaliana mR |
| 9b_bal2:BBFAF000270 | + 29.00 | 103.02 | 1.1e+03 | 2707 | I AF000270 Borrella burgdoferi |
| 9b_p11:GMEN0240 | - 29.00 | 102.30 | 1.1e+03 | 2960 | I X86442 G.max ENOD40-2 gene, 7/ |
| 9b_p11:CA255 | - 29.00 | 101.10 | 1.3e+03 | 3442 | I X70659 C.caibicans 25S rRNA, 2/ |
| 9b_p11:SP25SRNA | - 29.00 | 100.97 | 1.3e+03 | 3497 | I Z19136 S.pombe 25S rRNA gene. |

```

seq_name: gb_pr2:PTTRGC
seq_documentation_block: 166 bp mRNA PRI 21-MAY-1996
LOCUS PTTRGC
DEFINITION P.troglyodytes mRNA for TCR gamma constant chain.
ACCESSION X86723
VERSION X86723.1 GI:1568526
KEYWORDS TCR gamma chain constant region; TRGC gene.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Pan.
REFERENCE 1 (bases 1 to 166)
AUTHORS Zhang,X.M., Cathala,G., Soua,Z., Lefranc,M.P. and Huck.S.
TITLE The human T-cell receptor gamma variable pseudogene V10 is a
distinctive marker of human speciation
JOURNAL Immunogenetics 43 (4), 196-203 (1996)
MEDLINE 96163019
REFERENCE 2 (bases 1 to 166)
AUTHORS Lefranc,M.
TITLE Direct Submission
JOURNAL Submitted (26-APR-1995) M. Lefranc, Lab d'Immunogenetique
Moleculaire IIGM, Inst de Genetique Moleculaire, UMR 9942, BP 5051,
1919 Route de Mende, F-34033 Montpellier cedex 1, FRANCE
FEATURES
Source 1..166
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/rearranged
1..166
/gene="TRGC"
1..166
/gene="TRGC"
/segment="C-segment"
/product="TCR gamma constant chain"
53 a 37 c 28 g 48 t
BASE COUNT
ORIGIN
alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-5 x PTTRGC ..
Align seg 1/1 to: PTTRGC from: 1 to: 166
1 ArgilleLeuLeuArgTyr 6
|||||
116 CGGATATTATTAGATAC 133
seq_name: gb_p12:AF049680
seq_documentation_block: 204 bp DNA PLN 10-MAR-1998
LOCUS AF049680
DEFINITION Neotyphodium lolii strain Lp5 large subunit ribosomal RNA gene,
partial sequence.
ACCESSION AF049680
VERSION AF049680.1 GI:2944414
KEYWORDS
SOURCE Neotyphodium lolii.
ORGANISM Neotyphodium lolii.
Eukaryota; Fungi; Ascomycota; Pyrenomycetes; Hypocreales;
Clavicipitaceae; anamorphic Clavicipitaceae; Neotyphodium.
REFERENCE 1 (bases 1 to 204)
AUTHORS Ganley,A.R.D. and Scott,D.B.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1998) Institute of Molecular and Bio Sciences,
Massey University, Private Bag 11222, Palmerston North, New Zealand
FEATURES
Location/Qualifiers

```

```

source 1..204
/organism="Neotyphodium lolii"
/strain="Lp5"
/db_xref="taxon:73839"
/note="asexual form; PCR product from ntsl-nts2"
<1..>204
/note="from 3' region"
/product="large subunit ribosomal RNA"
46 a 45 c 66 g 47 t
BASE COUNT
ORIGIN
alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-5 x AF049680/rev ..
Align seg 1/1 to reverse of: AF049680 from: 1 to: 204
1 ArgilleLeuLeuArgTyr 6
|||||
29 CGAATTCGCTTCGGTAT 12
seq_name: gb_p12:AF049679
seq_documentation_block: 210 bp DNA PLN 10-MAR-1998
LOCUS AF049679
DEFINITION Epichloe typhina x Neotyphodium lolii strain Lp1 large subunit
ribosomal RNA gene, partial sequence.
ACCESSION AF049679
VERSION AF049679.1 GI:2944413
KEYWORDS
SOURCE Epichloe typhina x Neotyphodium lolii.
ORGANISM Epichloe typhina x Neotyphodium lolii
Eukaryota; Fungi; Ascomycota; Pyrenomycetes; Hypocreales;
Clavicipitaceae; Epichloe.
REFERENCE 1 (bases 1 to 210)
AUTHORS Ganley,A.R.D. and Scott,D.B.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1998) Institute of Molecular and Bio Sciences,
Massey University, Private Bag 11222, Palmerston North, New Zealand
FEATURES
Location/Qualifiers
1..210
/organism="Epichloe typhina x Neotyphodium lolii"
/strain="Lp1"
/db_xref="taxon:71426"
/note="PCR product from ntsl-nts2"
<1..>210
/note="from 3' region"
/product="large subunit ribosomal RNA"
47 a 48 c 68 g 47 t
BASE COUNT
ORIGIN
alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-5 x AF049679/rev ..
Align seg 1/1 to reverse of: AF049679 from: 1 to: 210
1 ArgilleLeuLeuArgTyr 6
|||||
32 CGAATTCGCTTCGGTAT 15
seq_name: gb_p12:AF049681

```

```

seq_documentation_block: 210 bp DNA PLN 10-MAR-1998
LOCUS AF049681
DEFINITION Epichloe typhina strain E8 large subunit ribosomal RNA gene,
partial sequence.
ACCESSION AF049681
VERSION AF049681.1 GI:2944415
KEYWORDS
SOURCE
ORGANISM
Epichloe typhina.
Epichloe typhina
Eukaryota; Fungi; Ascomycota; Pyrenomycetes; Hypocreales;
Clavicipitaceae; Epichloe.
REFERENCE
1 (bases 1 to 210)
AUTHORS Ganley,A.R.D. and Scott,D.B.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1998) Institute of Molecular and Bio Sciences,
Massey University, Private Bag 11222, Palmerston North, New Zealand
FEATURES
source
1..210
location/Qualifiers
/organism="Epichloe typhina"
/strain="E8"
/db_xref="taxon:5113"
/note="sexual form; PCR product from nts1-nts2"
<1..>210
/note="from 3' region"
/product="large subunit ribosomal RNA"
BASE COUNT 47 a 48 c 68 g 47 t
ORIGIN
1
|||||
264 CGAATTCTGCTCGGTAT 15

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-5 x AF049681/rev ..
Align seg 1/1 to reverse of: AF049681 from: 1 to: 210

1 ArgilleLeuArgTyr 6
|||||
32 CGAATTCTGCTCGGTAT 15

seq_name: gb_pr1: HSAJ8203

seq_documentation_block: 300 bp mRNA PRI 29-APR-1998
LOCUS HSAJ8203
DEFINITION Homo sapiens mRNA for immunoglobulin heavy chain, VHDJH
rearrangement : VHJ146.
ACCESSION AJ008203
VERSION AJ008203.1 GI:3097118
KEYWORDS igh; immunoglobulin heavy chain.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 300)
AUTHORS Jahn,S.
TITLE Direct Submission
JOURNAL Submitted (26-APR-1998) Jahn S., Department of Dermatology, Medical
Faculty (Charite), Humboldt University Berlin, Schumannstr. 20/21,
10117 Berlin, GERMANY
REFERENCE
2 (bases 1 to 300)
AUTHORS Gellrich,S., Rutz,S., Borkowski,A., Golembowski,S.,
Gronmich-Inne,E., Steery,W. and Jahn,S.
TITLE Analysis of VHDJH gene transcripts in B cells infiltrating salivary
glands and lymph node tissue in patients with Sjogren's syndrome
JOURNAL Unpublished
FEATURES
source
1..300
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="14"

seq_documentation_block: 360 bp mRNA PRI 02-JUN-1999
LOCUS HSA244960
DEFINITION Homo sapiens mRNA for immunoglobulin mu heavy chain variable
region, partial, clone 1-B73.
ACCESSION AJ244960
VERSION AJ244960.1 GI:4995379
KEYWORDS igh; IgM heavy chain; immunoglobulin mu heavy chain; variable
region.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 360)
AUTHORS Dono,M., Zupo,S., Chiorazzi,N. and Ferrarini,M.
TITLE Heterogeneity of tonsillar subepithelial B lymphocytes, the splenic
marginal zone equivalents
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 360)
AUTHORS Dono,M.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-1999) Dono M., Istituto Nazionale per la Ricerca
sul Cancro, Immunologia Clinica, L.go Rosanna benzi 10- Genova,
ITALY
FEATURES
source
1..360
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="14"
/rearranged
/tissue_type="tonsil"
/cell_type="B-lymphocyte"
/clone="1-B73"
<1..>360
/gene="IGHV4-34"
/product="immunoglobulin mu heavy chain variable region"
1..360
/gene="IGHV4-34"
<1..>360
/gene="IGHV4-34"
/codon_start=1
/product="immunoglobulin mu heavy chain variable region"
/db_xref="GI:4995380"

```

```

/sex="Female"
/cell_type="B lymphocyte"
/dev_stage="adult"
/map="q32.33"
/tissue_type="lymph node"
<1..>300
/gene="VHDJH"
/product="immunoglobulin heavy chain"
1..300
/gene="VHDJH"
BASE COUNT 67 a 78 c 87 g 68 t
ORIGIN
1
|||||
264 CGAATTCTAGCTAT 281

seq_name: gb_pr2: HSA244960

seq_documentation_block: 360 bp mRNA PRI 02-JUN-1999
LOCUS HSA244960
DEFINITION Homo sapiens mRNA for immunoglobulin mu heavy chain variable
region, partial, clone 1-B73.
ACCESSION AJ244960
VERSION AJ244960.1 GI:4995379
KEYWORDS igh; IgM heavy chain; immunoglobulin mu heavy chain; variable
region.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 360)
AUTHORS Dono,M., Zupo,S., Chiorazzi,N. and Ferrarini,M.
TITLE Heterogeneity of tonsillar subepithelial B lymphocytes, the splenic
marginal zone equivalents
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 360)
AUTHORS Dono,M.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-1999) Dono M., Istituto Nazionale per la Ricerca
sul Cancro, Immunologia Clinica, L.go Rosanna benzi 10- Genova,
ITALY
FEATURES
source
1..360
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="14"
/rearranged
/tissue_type="tonsil"
/cell_type="B-lymphocyte"
/clone="1-B73"
<1..>360
/gene="IGHV4-34"
/product="immunoglobulin mu heavy chain variable region"
1..360
/gene="IGHV4-34"
<1..>360
/gene="IGHV4-34"
/codon_start=1
/product="immunoglobulin mu heavy chain variable region"
/db_xref="GI:4995380"

```

/translation="OVLOQWAGLLKPSSESLTCAVYGGSLGYVWSWIROPKKG
LENGEINYSGSNYPKSLASRVITISVDTSKQFSLKLTSTADTAIVYWCARGNEAG
YYGMDVGGGTWTVSS"

BASE COUNT 77 a 95 c 113 g 75 t

ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-5 x HSA244960 ..

Align seg 1/1 to: HSA244960 from: 1 to: 360

1 ArgilleLeuLeuArgTyr 6

|||||
303 CGGATATTACTACGGTAT 320

seq_name: gb_sts:G14238

seq_documentation_block:

LOCUS G14238 400 bp DNA STS 22-DEC-1995
DEFINITION human STS SHGC-7374 clone pg-3478.

ACCESSION G14238

VERSION G14238.1 GI:1129977

KEYWORDS STS sequence; primer; sequence tagged site.

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Eukaryotes; Metazoa; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
Carnivora; Hominidae; Homo.

REFERENCE

AUTHORS

JOURNAL

COMMENT

Contact: Richard M. Myers

Stanford Human Genome Center (SHGC)

Stanford University School of Medicine

Department of Genetics, M-344, Stanford, CA 94305, USA

Tel: 4157259687

Fax: 4157259689

Email: myers@shgc.stanford.edu

Primer A: GCAATGCCAGATTGGAATAA

Primer B: AGGCCTATGCTCCCTCAATAA

STS size: 152

PCR Profile:

Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds

Annealing: 62 degrees C for 23 seconds

Polymerization: 72 degrees C for 30 seconds

PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9600

Protocol:

Template: 25 ng

Primer: each 1 uM

dNTPs: each 200 uM

Tag Polymerase: 0.05 units/ul

Total Vol: 10 ul

Buffer:

MgCl2: 2.5 mM

KCl: 50 mM

Tris-HCl: 20 mM

pH: 8.3

Plasmid clones, generated from a lymphoblastoid cell line from a human male. Localized to human chromosome 17 by analysis on the

NIGMS Human/Rodent Somatic Cell Hybrid Panel #1, Coriell Institute
for Medical Research, Camden, NJ 08103.

FEATURES

source

1..400

/organism="Homo sapiens"

STS

186..337

primer_bind

186..206

primer_bind

complement(317..337)

BASE COUNT 115 a 83 c 88 g 111 t 3 others

ORIGIN

alignment_scores:

Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-5 x G14238/rev ..

Align seg 1/1 to reverse of: G14238 from: 1 to: 400

1 ArgilleLeuLeuArgTyr 6

|||||

388 AGAATCCTCTTGAGGTAT 371

seq_name: gb_sts:KLAJ9903

seq_documentation_block:

LOCUS KLAJ9903 436 bp DNA STS 20-NOV-1998

DEFINITION Kluyveromyces lactis DNA fragment for sequence tagged site, clone

ZZ.

ACCESSION AJ229903

VERSION AJ229903.1 GI:3820353

KEYWORDS STS.

SOURCE Kluyveromyces lactis.

ORGANISM Kluyveromyces lactis.

Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;

Saccharomycetaceae; Kluyveromyces.

REFERENCE 1 (bases 1 to 436)

AUTHORS Ozier-Kalogeropoulos,O.

TITLE Direct Submission

JOURNAL Submitted (28-MAY-1998) Ozier-Kalogeropoulos O., Institut Pasteur,

Genetique Moleculaire des Levures25 rue de Dr Roux, 75724 Paris

Cedex 15 FRANCE

REFERENCE 2 (bases 1 to 436)

AUTHORS Ozier-Kalogeropoulos,O., Malpertuy,A., Boyer,J., Tekala,F. and

Dujon,B.

TITLE Random exploration of the Kluyveromyces lactis genome and

comparison with that of Saccharomyces cerevisiae

JOURNAL Nucleic Acids Res. 26 (23), 5511-5524 (1998)

MEDLINE 99045620

FEATURES

source

1..436

/organism="Kluyveromyces lactis"

/strain="CBS2359"

/db_xref="taxon:28985"

/clone="ZZ"

STS 113 a 86 c 110 g 127 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-5 x KLAJ9903/rev ..

Align seg 1/1 to reverse of: KLAJ9903 from: 1 to: 436

1 ArgilleLeuArgTyr 6
|||||
147 CGAATCTGCTCGGTAT 130

seq_name: gb_pr3:AF062148

seq_documentation_block:

LOCUS AF062148 441 bp mRNA PRI 02-JUN-1998
DEFINITION Homo sapiens clone 45u-23 immunoglobulin heavy chain variable region (IGH) mRNA, partial cds.

ACCESSION AF062148
VERSION AF062148.1 GI:3170758
KEYWORDS
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 441)
Wang,X. and Stollar,B.D.
Autoreactivity and immunoglobulin VH gene expression in aging humans

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 441)

Wang,X. and Stollar,B.D.

Direct Submission

TITLE

Submitted (22-APR-1998) Biochemistry Department, Tufts University

School of Medicine, 136 Harrison Ave., Boston, MA 02111, USA

FEATURES

Location/Qualifiers

1..441

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="14"

/map="14q32.33"

/clone="45u-23"

/cell_type="peripheral B lymphocyte"

/tissue_type="blood"

/note="from elderly repertoire 45u"

1..>441

/gene="IGH"

1..>441

/gene="IGH"

/codon_start=1

/product="immunoglobulin heavy chain variable region"

/protein_id="AAC18184.1"

/db_xref="GI:3170758"

/translation="MKHLWFLLLVAAPRLVLSQVQLQESGPGLVKPSQTLSTCTVVS

GGSSISGGYVSWIRHPGKLEWIGIYIGSTYINPSLKRVTISVDFSKNQFSLK

LSSVTADTAVYICARDHVVVVAPEYIYGMVYGQTTVTVSSG"

1..57

/gene="IGH"

58..>441

/gene="IGH"

89 a 126 c 127 g 99 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 29.00

Ratio: 4.833

Percent Similarity: 100.000

Length: 6

Gaps: 0

Percent Identity: 100.000

alignment_block:

US-08-653-294-5 x AF062148

Align seg 1/1 to: AF062148 from: 1 to: 441

1 ArgilleLeuArgTyr 6

|||||

381 CGAATCTACTACGGTAT 398

seq_name: gb_pr3:AF056445

seq_documentation_block:

LOCUS AF056445 443 bp mRNA PRI 16-MAY-1998
DEFINITION Homo sapiens clone TCB1 Cri-du-chat critical region mRNA.
ACCESSION AF056445
VERSION AF056445.1 GI:3044163
KEYWORDS
SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 443)

Simmons,A.D. and Lovett,M.L.

High resolution physical and transcription maps of the Cri-du-chat critical region

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 443)

Simmons,A.D. and Lovett,M.L.

Direct Submission

TITLE

Submitted (31-MAR-1998) Department of Otorhinolaryngology,

Molecular Biology and Oncology, The McDermott Center, University of

Texas Southwestern Medical Center, 5323 Harry Hines Boulevard,

Dallas, TX 75235-8591, USA

FEATURES

Location/Qualifiers

1..443

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="5"

/map="5p15.2"

/clone="TCB1"

1..443

/note="Cri-du-chat critical region"

BASE COUNT 108 a 143 c 55 g 137 t

ORIGIN

alignment_scores:

Quality: 29.00

Ratio: 4.833

Percent Similarity: 100.000

Length: 6

Gaps: 0

Percent Identity: 100.000

alignment_block:

US-08-653-294-5 x AF056445/rev

Align seg 1/1 to reverse of: AF056445 from: 1 to: 443

1 ArgilleLeuArgTyr 6

|||||

231 AGGATATTGTAAGTAC 214

seq_name: gb_sts:G23069

seq_documentation_block:

LOCUS G23069 455 bp DNA STS 31-MAY-1996

DEFINITION human STS WI-12710, sequence tagged site.

ACCESSION G23069

VERSION G23069.1 GI:1343395

KEYWORDS STS; STS sequence; primer; sequence tagged site.

SOURCE human STSs derived from sequences in dbEST and the Unigene collection.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 455)

Hudson,T.

Whitehead Institute/MIT Center for Genome Research; Physically

Mapped STSs

Unpublished (1995)

JOURNAL

COMMENT

Contact: Thomas Hudson

Whitehead Institute/MIT Center for Genome Research

Whitehead Institute for Biomedical Research

9 Cambridge Center, Cambridge MA 02142 USA

Tel: 617 252 1900

Fax: 617 252 1902

Email: thudson@genome.wi.mit.edu

Primer A: CAGTTTGTAGTGACATGTGAGG
Primer B: AACAGCTGCTGCCCTCAG
STS size: 127
PCR Profile:

Presoak:
Denaturation:
Annealing: 56 degrees C
Polymerization:
PCR Cycles: 35
Thermal Cycler:
Protocol:
Template: 10 ng
Primer: each 5 pM
dNTPs: each 4 nM
Taq Polymerase: 0.025 units/ul
Total Vol: 20 ul

Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCL: 10 mM
pH: 9.3

Derived from dbEST (genbank accession H49567).

FEATURES
source
Location/Qualifiers
1..455
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="750.8 CR from top of Chr1 linkage group"
19..145
19..142
primer_bind
primer_bind
BASE COUNT 104 a 106 c 118 g 123 t 4 others
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-5 x G23069/rev ..
Align seg 1/1 to reverse of: G23069 from: 1 to: 455

1 ArgilleLeuLeuArgTyr 6
321 AGAATCTTACTCAGGTAT 304

seq_name: gb_sts:KLAJ9696

seq_documentation_block:
LOCUS KLAJ9696 455 bp DNA STS 20-NOV-1998
DEFINITION Kluyveromyces lactis DNA fragment for sequence tagged site, clone 22.

ACCESSION AJ229696
VERSION AJ229696.1 GI:3820146
KEYWORDS STS.
SOURCE Kluyveromyces lactis.
ORGANISM Kluyveromyces lactis

REFERENCE 1 (bases 1 to 455)
AUTHORS Ozier-Kalogeropoulos O.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-1998) Ozier-Kalogeropoulos O., Institut Pasteur, Genetique Moleculaire des Levures 25 rue de Dr Roux, 75724 Paris cedex 15 FRANCE
REFERENCE 2 (bases 1 to 455)
AUTHORS Ozier-Kalogeropoulos O., Malpertuy A., Boyer J., Tekala F. and

Dujon B.
Random exploration of the Kluyveromyces lactis genome and comparison with that of Saccharomyces cerevisiae
Nucleic Acids Res. 26 (23), 5511-5524 (1998)
99045620
FEATURES
source
Location/Qualifiers
1..455
/organism="Kluyveromyces lactis"
/strain="CBS2359"
/db_xref="taxon:28985"
/clone="22"
/c1..>455

BASE COUNT 129 a 115 c 96 g 115 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-5 x KLAJ9696 ..
Align seg 1/1 to: KLAJ9696 from: 1 to: 455

1 ArgilleLeuLeuArgTyr 6
354 CGAATCTGCTCGGTAT 371

seq_name: gb_pl1:HAU56781

seq_documentation_block:
LOCUS HAU56781 588 bp DNA PLN 01-NOV-1996
DEFINITION Hampea appendiculata 5.8S ribosomal RNA gene and internal transcribed spacer 1 and 2.

ACCESSION U56781
VERSION U56781.1 GI:1654304
KEYWORDS Hampea appendiculata.
SOURCE Hampea appendiculata

ORGANISM Hampea appendiculata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Hampea.

REFERENCE 1 (bases 1 to 588)
AUTHORS Seelanan, T., Wendel, J.F. and Schnabel, A.
TITLE Congruence and consensus in the cotton tribe: Evidence from the nuclear and plastid genomes
Syst. Bot. (1996) In press

REFERENCE 2 (bases 1 to 588)
AUTHORS Seelanan, T., Wendel, J.F. and Schnabel, A.
TITLE Direct Submission
JOURNAL Submitted (25-APR-1996) T. Seelanan, Botany, Iowa State University, Ames, IA 50011, USA

FEATURES
source
Location/Qualifiers
1..588
/organism="Hampea appendiculata"
/db_xref="taxon:47613"

misc_RNA
1..218
/note="internal transcribed spacer 1"
219..382
/note="5.8S ribosomal RNA"

misc_RNA
383..588
/note="internal transcribed spacer 2"
BASE COUNT 146 a 158 c 145 g 139 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```
alignment_block:
US-08-653-294-5 x HAU56781/rev
Align seg 1/1 to reverse of: HAU56781 from: 1 to: 588

1 ArgileLeuLeuArgTyr 6
|||||
280 CGCATTTTGCTACGTTAT 263

seq_name: gb_pl2:CNS01995

seq_documentation_block:
LOCUS      CNS01995      636 bp      mRNA      PLN      02-SEP-1999
DEFINITION Botrytis cinerea strain T4 cDNA library under conditions of
            nitrogen deprivation.
ACCESSION  AL111345
VERSION    AL111345.1 GI:5825964
KEYWORDS   cDNA library; nitrogen deprivation.
SOURCE     Botryotinia fuckeliana.
ORGANISM   Botryotinia fuckeliana
            Eukaryota; Fungi; Ascomycota; Euascomycetes; Discomycetes;
            Leotiales; Sclerotiniaceae; Botryotinia.
REFERENCE  1 (bases 1 to 636)
AUTHORS    Bitton, F., Levis, C., Fortini, D., Pradier, J.M. and Brygoo, Y.
TITLE      Direct Submission
JOURNAL    Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,
            78026 Versailles, France
REFERENCE  2 (bases 1 to 636)
AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage :
            CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT    The cDNA library to be analyzed within the framework of this
            project was created using a Botrytis cinerea strain which was grown
            under conditions of nitrogen deprivation, which is the normal
            situation for B. cinerea during its development on its host plant.
            The library was produced in an oriented direction, in the pBSII
            vector.

FEATURES             Location/Qualifiers
     source            1..636
                     /organism="Botryotinia fuckeliana"
                     /strain="T4"
                     /db_xref="taxon:40559"
                     /notes="Genoscope sequence ID : W69H121"
BASE COUNT  162 a 133 c 172 g 168 t 1 others
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-5 x CNS01995/rev
Align seg 1/1 to reverse of: CNS01995 from: 1 to: 636

1 ArgileLeuLeuArgTyr 6
|||||
454 CGAATTCGCTTCGGTAT 437
```

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-5 to: N_Geneseq_36.* out_format : pfs
 Date: Feb 8, 2000 1:27 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
 -O=/cgnl_1/USPTO_spool/US08653294/runat_04022000_160701_15807/app_query.fasta.1
 -DB=N_Geneseq_36 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOFLC=0.000 -LOOPEXT=0.000
 -GAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -GAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT
 -THREADS=1

Search information block:

Query: US-08-653-294-5
 Query length: 6
 Database: N_Geneseq_36.*
 Database sequences: 311585
 Database length: 125096042
 Search time (sec): 590.520000

score_list:

| Sequence | Strd | Orig | Zscore | EScore | Len | Documentation |
|------------------------|------|-------|--------|---------|--------|----------------------------------|
| N_Geneseq_36:V74992 | - | 29.00 | 108.71 | 186.73 | 863 | Staphylococcus aureus contig SE |
| N_Geneseq_36:V34296 | - | 29.00 | 103.16 | 380.64 | 1688 | Human secreted protein gene 11 |
| N_Geneseq_36:X20304 | + | 29.00 | 101.26 | 485.54 | 2223 | Borrelia burgdorferi polynucle |
| N_Geneseq_36:Q74809 | + | 29.00 | 100.88 | 509.85 | 2123 | Aspartokinase II gene. DNA enc |
| N_Geneseq_36:V03795 | + | 29.00 | 96.93 | 845.79 | 3591 | PAMV coat protein coding sequ |
| N_Geneseq_36:X13280 | + | 29.00 | 96.21 | 927.75 | 3907 | Enterococcus faecalis genome o |
| N_Geneseq_36:V36211 | - | 29.00 | 95.50 | 1.0e+03 | 4256 | Total contiguous sequence of E |
| N_Geneseq_36:V83003 | - | 29.00 | 95.50 | 1.0e+03 | 4256 | Contiguous sequence determined |
| N_Geneseq_36:X12982 | - | 29.00 | 84.75 | 4.0e+03 | 15614 | Enterococcus faecalis genome |
| N_Geneseq_36:Q26060 | - | 28.00 | 127.88 | 15.98 | 51 | SERP gene PCR primer p3. New pla |
| N_Geneseq_36:T26893 | - | 28.00 | 107.38 | 221.51 | 607 | Haemophilus influenzae 15 kd ou |
| N_Geneseq_36:Q42561 | - | 28.00 | 94.27 | 1.2e+03 | 2960 | Histamine H1 receptor coding s |
| N_Geneseq_36:T2653 | + | 28.00 | 90.59 | 1.9e+03 | 4617 | Rat tripeptidylpeptidase II cD |
| N_Geneseq_36:T71315 | - | 28.00 | 89.81 | 2.1e+03 | 5072 | CDNA encoding C. elegans UNC-5 |
| N_Geneseq_36:T71314 | - | 28.00 | 89.81 | 2.1e+03 | 5073 | CDNA encoding C. elegans UNC-5 |
| N_Geneseq_36:T72654 | + | 28.00 | 89.75 | 2.1e+03 | 5109 | Rat brain homologue of serine |
| N_Geneseq_36:T71320 | - | 28.00 | 86.60 | 3.2e+03 | 7474 | Plasmid pTB73 encoding C. eleg |
| N_Geneseq_36:T71324 | - | 28.00 | 84.49 | 4.2e+03 | 9642 | Plasmid pTB113 used for over-e |
| N_Geneseq_36:T71319 | - | 28.00 | 83.83 | 4.5e+03 | 10443 | Plasmid pTB72 encoding C. eleg |
| N_Geneseq_36:T71321 | - | 28.00 | 81.76 | 5.9e+03 | 13414 | Plasmid pCB50 encoding UNC-53 |
| N_Geneseq_36:V12029_05 | - | 28.00 | 84.34 | 5.3e+04 | 110000 | Continuation (6 of 17) of |
| N_Geneseq_36:T21689 | + | 27.00 | 126.99 | 17.91 | 34 | Acromonium cellulolyticus cellul |
| N_Geneseq_36:Q55183 | - | 27.00 | 97.50 | 786.43 | 1199 | Sequence for a synthetic lacP- |
| N_Geneseq_36:Q62660 | - | 27.00 | 97.50 | 786.43 | 1199 | gp41 HIV envelope protein codi |
| N_Geneseq_36:N90797 | - | 27.00 | 97.47 | 789.91 | 1204 | Synthetic p41 gene with p120 c |
| N_Geneseq_36:V74890 | + | 27.00 | 96.91 | 847.84 | 1287 | Staphylococcus aureus contig S |
| N_Geneseq_36:V74636 | + | 27.00 | 96.28 | 919.34 | 1389 | Staphylococcus aureus contig S |
| N_Geneseq_36:V30591 | - | 27.00 | 95.82 | 975.65 | 1469 | Clostridium botulinum type B t |
| N_Geneseq_36:V22682 | - | 27.00 | 92.42 | 1.5e+03 | 2214 | New DNA sequence isolated from |
| N_Geneseq_36:T63530 | + | 27.00 | 90.91 | 1.8e+03 | 2659 | Pichia pastoris HFS4 gene. Pol |
| N_Geneseq_36:Q23074 | + | 27.00 | 90.25 | 2.0e+03 | 2879 | Bacterial alpha-1,3-glucan 3-g |
| N_Geneseq_36:V22683 | - | 27.00 | 89.04 | 2.3e+03 | 3331 | New DNA sequence isolated from |
| N_Geneseq_36:X52264 | - | 27.00 | 88.26 | 2.6e+03 | 3662 | Protein PRO335 cDNA clone DNA |
| N_Geneseq_36:Q44362 | - | 27.00 | 87.98 | 2.7e+03 | 3789 | Sequence of bovine ephemeral f |
| N_Geneseq_36:X52266 | - | 27.00 | 87.42 | 2.9e+03 | 4053 | Protein PRO326 cDNA clone DNA |
| N_Geneseq_36:X23315 | + | 27.00 | 86.13 | 3.4e+03 | 4736 | Mouse 1-alpha-OHase promoter r |
| N_Geneseq_36:V12368 | + | 27.00 | 84.13 | 4.4e+03 | 6032 | Trichoderma harzianum mutanase |
| N_Geneseq_36:V30268 | + | 27.00 | 84.13 | 4.4e+03 | 6032 | Plasmid pMT1802 encoding Trich |
| N_Geneseq_36:T35165 | - | 27.00 | 81.63 | 6.0e+03 | 8157 | Plasmid pPHIL-B2 used for expre |
| N_Geneseq_36:T35166 | - | 27.00 | 81.21 | 6.3e+03 | 8584 | Plasmid pPHIL-B2 (MfalphaprePro |
| N_Geneseq_36:T35168 | - | 27.00 | 81.20 | 6.3e+03 | 8590 | Plasmid pD2p1ck(MfalphaprePro |

N_Geneseq_36:V28852 - 27.00 81.19 6.3e+03 8598 1 pPIC9/ELF251. cDNA construct
 N_Geneseq_36:T77819 - 27.00 80.84 6.6e+03 8974 1 Neisseria meningitidis clas
 N_Geneseq_36:T77817 - 27.00 80.67 6.8e+03 9156 1 Neisseria meningitidis clas
 N_Geneseq_36:T77818 - 27.00 80.64 6.8e+03 9191 1 Neisseria meningitidis clas

seq_name: N_Geneseq_36:V74992

seq_documentation_block:

ID V74992 standard; DNA; 863 BP.
 AC V74992;
 DT 16-MAR-1999 (first entry)
 DE Staphylococcus aureus contig SEQ ID #681.
 KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.
 OS Staphylococcus aureus;
 FH Key Location/Qualifiers
 misc_feature 181..240
 FT /*tag= a
 FT /*note= "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"

EP-786519-A2.

30-JUL-1997.

07-JAN-1997; 100117

PR 05-JAN-1996; US-009861.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,

PI Rosen CA;

DR WPI: 97-374922/35.

PT Polynucleotide(s) and proteins derived from Staphylococcus aureus
 PT stored on computer readable medium and used in the production of
 PT anti-S.aureus vaccines

PS Claim 1; Page 1600-1601; 3271pp; English.

CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S.aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against S.aureus infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S.aureus DNA sequences contained on the
 CC computer readable medium.

SQ Sequence 863 BP; 252 A; 171 C; 96 G; 280 T;

alignment_scores:

Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-5 x V74992/rev ..

Align seg 1/1 to reverse of: V74992 from: 1 to: 863

1 ArgilleleuArgTyr 6

146 AGGATACTCTTAGAT 129

seq_name: N_Geneseq_36:V34296

seq_documentation_block:

ID V34296 standard; DNA; 1688 BP.
 AC V34296;
 DT 29-JAN-1999 (first entry)
 DE Human secreted protein gene 11 clone HLTBS22.
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 OS Homo sapiens.
 PN WO9804083-A2.
 PD 17-SEP-1998.
 PF 12-MAR-1998; U04858.
 PR 19-DEC-1997; US-068368.
 PR 14-MAR-1997; US-040710.
 PR 14-MAR-1997; US-040762.
 PR 30-MAY-1997; US-048100.
 PR 30-MAY-1997; US-048189.
 PR 30-MAY-1997; US-048357.
 PR 06-JUN-1997; US-050934.
 PR 05-SEP-1997; US-048970.
 PR 05-SEP-1997; US-057765.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ferris AM, Fischer CL, Gentz RL, Greene JM, Kyaw H,
 PI Li H, Li Y, Moore PA, Rosen CA, Ruben SM, Soppet DR,
 PI Wei YF, Young PE, Zeng Z;
 DR WPI: 98-520811/44.
 DR P-PSDB; W75206.

PT Isolated human poly:nucleotide(s) encoding secretory peptide(s) -
 used to develop products for the diagnosis and treatment of e.g.
 PT inflammation, cancers, CNS disorders or immune system disorders
 PS Claim 1: Page 123-124; 201pp; English.
 CC This sequence represents a nucleic acid molecule which encodes a secreted
 CC human protein. The gene number, and the clone it is derived from, are
 CC detailed in the descriptor line. The gene can be used to generate fusion
 CC proteins by linking to the gene to a human immunoglobulin Fc portion
 CC (e.g. V34277) for increasing the stability of the fused protein as
 CC compared to the human protein only.
 CC The invention relates to 28 novel genes and their fragments (nucleic acid
 CC sequences: V34286-V34325; amino acid sequences W75196-W75235) which
 CC are useful for preventing, treating or ameliorating medical conditions
 CC e.g. by protein or gene therapy. Also, pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new polynucleotides.
 CC Specific uses are described for each of the 28 polynucleotides, based on
 CC which tissues they are most highly expressed in (see V34286 for described
 CC uses).
 SQ Sequence 1688 BP; 443 A; 430 C; 387 G; 427 T;

alignment_scores:

Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-5 x V34296/rev ..

Align seg 1/1 to reverse of: V34296 from: 1 to: 1688

1 ArgilleLeuArgTyr 6

|||||
 394 AGGATCCCTTCGAGGTAT 377

seq_name: N_Geneseq_36:X20304

seq_documentation_block:

ID X20304 standard; DNA; 2123 BP.
 AC X20304;

DT 04-MAY-1999 (first entry)
 DE Borrelia burgdorferi polynucleotide sequence #57.
 KW Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
 KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
 KW infection; diagnosis; characterisation; detection; ds.
 OS Borrelia burgdorferi.
 PN WO9858943-A1.
 PD 30-DEC-1998.
 PF 18-JUN-1998; U12764.
 PR 03-SEP-1997; US-057483.
 PR 20-JUN-1997; US-050359.
 PR 22-JUL-1997; US-053344.
 PR 22-JUL-1997; US-053377.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (MEDI-) MEDIMUNE INC.
 PI Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO,
 PI White OR;
 DR WPI: 99-081217/07.
 PT New isolated Borrelia burgdorferi nucleic acids - used to develop
 PT products for the detection, diagnosis, characterisation, prevention
 PT and therapy of infections, particularly Lyme disease
 PS Claim 1: Page 1037-1039; 1128pp; English.
 CC X20248 to X20402 represent polynucleotide sequences isolated from
 CC Borrelia burgdorferi (Bb). Products derived from Bb can be used for
 CC the detection, diagnosis, characterisation, prevention and therapy of
 CC Bb infections, e.g. Lyme disease. They can also be used for the
 CC production of biosynthetic products, e.g. enzymes. Borrelia belongs
 CC to a family of motile, spiral-shaped bacteria called Spirochetes.
 CC Spirochetes are pathogenic in humans and Borrelia causes epidemic and
 CC endemic relapsing fever, and Lyme borreliosis, more commonly known as
 CC Lyme disease.
 SQ Sequence 2123 BP; 765 A; 327 C; 262 G; 763 T;

alignment_scores:

Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-5 x X20304 ..

Align seg 1/1 to: X20304 from: 1 to: 2123

1 ArgilleLeuArgTyr 6

|||||
 500 CGTATCTTTTAAAGATAC 517

seq_name: N_Geneseq_36:Q47809

seq_documentation_block:

ID Q47809 standard; DNA; 2223 BP.
 AC Q47809;

DT 10-MAR-1994 (first entry)

DE Aspartokinase II gene.

Dimer: subunit; lysine-sensitive; aspartokinase II; AKII;

Bacillus sp.; MGA3; methylotrophic; thermotolerant; tetramer;

complementation; E. coli; auxotrophic mutant; isozyme; lysine; ss.

OS Bacillus sp. MGA3.

FH Key Location/Qualifiers

FT -35_signal 259..264

FT /*tag= a

FT -10_signal 283..288

FT /*tag= b

FT misc_signal 297

FT /*tag= c

FT repeat_unit 490..513

FT /*tag= d

FT repeat_unit 526..576

FT /*tag= e

FT /*rpt_type= INVERTED

```

FT /note= "Indicated in the spec. as having dyad symmetry
FT to the region 490-513"
FT cds 664..1899
FT /tag= f
FT /product= Lysine-sensitive aspartokinase II
FT misc_difference 757..759
FT /tag= g
FT /codon= seq:ACC aa:Asn
FT misc_difference 1237..1239
FT /tag= h
FT /codon= seq:CCT aa:Leu
FT repeat_unit 1936..1980
FT /tag= i
FT /rpt_type= INVERTED
FT repeat_unit 1987..2028
FT /tag= j
FT /rpt_type= INVERTED
FT
PN US243039-A.
PD 07-SEP-1993.
PF 335691.
PR 10-APR-1989; US-335691.
PR 10-APR-1989; US-335691.
PR 12-MAY-1989; US-351436.
PR 20-MAR-1991; US-673263.
PR 20-MAR-1991; US-673264.
PR 12-APR-1991; US-684135.
PR (MINU ) UNIV MINNESOTA.
PA Flickinger MC, Schendel FJ;
DR WPI: 93-295300/37.
DR P-PSDB: R48046.
PT DNA encoding lysine-sensitive aspartokinase II - from
PT methylothrophic, thermotolerant bacillus sp. MGA 3, for producing
PT lysine-overproducing mutants
PS Claim 2: Fig 2: 16pp: English.
CC This sequence encodes a polypeptide which corresponds to the
CC alpha2 subunit of lysine-sensitive aspartokinase II (AKII)
CC of bacillus sp. MGA3. MGA3 is methylothrophic and thermotolerant.
CC The native form of AKII is an alpha2beta2 tetramer. This DNA
CC sequence was identified by cloning the structural gene from a
CC genomic library via complementation of an E. coli auxotrophic mutant
CC lacking all three aspartokinase isozymes. The open reading frame
CC shown encodes both the alpha and beta subunits in the same reading
CC frame. This gene may be used to produce microorganisms such as new
CC strains of bacteria which overproduce lysine at higher levels or
CC under more stringent environmental conditions.
SQ Sequence 2223 BP; 702 A; 394 C; 498 G; 629 T;

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-5 x Q47809 ..
Align seg 1/1 to: Q47809 from: 1 to: 2223

seq_name: N_Geneseq_36:V03795
seq_documentation_block:
ID V03795 standard; cDNA to mRNA; 3581 BP.
AC V03795;
DT 22-APR-1998 (first entry)
DE PAMMV coat protein coding sequence.
KW Coat protein; Patchouli mild mosaic virus; PAMMV; ss.
OS Patchouli mild mosaic virus.
FH Key Location/Qualifiers
FT CDS 1591..3390
FT /tag= a

```

```

FT misc_feature 1591..2796
FT /tag= b
FT /note= "encodes 42K coat protein (see W41020)"
FT misc_feature 2797..3387
FT /tag= c
FT /note= "encodes 26K coat protein (see W41021)"
FT
PN J09154582-A.
PD 17-JUN-1997.
PF 316431.
PR 05-DEC-1995; JP-316431.
PR 05-DEC-1995; JP-316431.
PA (KAOS ) KAO CORP.
DR WPI: 97-367062/34.
DR P-PSDB: W41020, W41021, W41022.
PT Patchouli mild mosaic virus coat protein gene - useful for
PT recombinant expression of viral protein
PS Example 5: Page 4-7; 10pp; Japanese.
CC This sequence represents the coding sequence for the coat protein of
CC patchouli mild mosaic virus (PAMMV), and is a DNA sequence of the
CC invention. The coat protein of PAMMV, which cannot be synthesised
CC naturally from a plant genome, can be produced by the plant cell
CC transformed by the DNA.
SQ Sequence 3581 BP; 1020 A; 654 C; 852 G; 1055 T;

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-5 x V03795 ..
Align seg 1/1 to: V03795 from: 1 to: 3581

seq_name: N_Geneseq_36:X13280
seq_documentation_block:
ID X13280 standard; DNA; 3907 BP.
AC X13280;
DT 19-MAR-1999 (first entry)
DE Enterococcus faecalis genome contig SEQ ID NO:343.
KW Enterococcus faecalis; contig; detection; Enterococcal infection;
KW vaccine; attenuation; computer readable medium; ds.
OS Enterococcus faecalis.
PN W09850555-A2.
PD 12-NOV-1998.
PF 04-MAY-1998; U08985.
PR 14-NOV-1997; US-066009.
PR 06-MAY-1997; US-044031.
PR 16-MAY-1997; US-046655.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Dillon PJ, Kunsch CA;
DR WPI: 99-045171/04.
PT New isolated Enterococcus faecalis polynucleotides and polypeptides
PT - used to develop products for the detection of Enterococcus and for
PT use in vaccines for prevention or attenuation of Enterococcus
PT infection.
PS Claim 1: Page 1461-1463; 2084pp; English.
CC A computer readable medium has been developed which has recorded on it
CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
CC X12938 to X13919 represent these nucleotide sequences which are primary
CC nucleotide sequences, also known as contigs. The computer-based system
CC can identify fragments of the Enterococcus faecalis genome with
CC commercial importance. The products can be used to detect the presence
CC of Enterococcus faecalis in samples. They can also be used for
CC diagnosing Enterococcal infection in an animal and monitoring
CC progression of disease, and for identifying agents which can be used to
CC modulate the growth or pathogenicity of Enterococcus faecalis, or
CC another related organism, in vivo or in vitro. In particular the

```

CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
 CC can be used in vaccines to prevent or attenuate an Enterococcal
 CC infection.
 SQ Sequence 3907 BP; 1056 A; 598 C; 717 G; 1424 T;

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-5 x X13280 ..

Align seg 1/1 to: X13280 from: 1 to: 3907

1 ArgilleLeuLeuArgTyr 6

|||||

968 AGAATATTGAGGTAT 985

seq_name: N_Geneseq_36.V36211

seq_documentation_block:

ID V36211 standard; DNA; 4256 BP.

AC V36211;

DT 03-SEP-1998 (first entry)

DE Total contiguous sequence of P. carinii in immunosuppressed rats.

KW Detection; diagnosis; 26S rRNA gene; P. carinii specific; infection;

KW Species identification; ss.

OS Pneumocystis carinii.

PN US5776680-A.

PD 07-JUL-1998.

PF 21-JUL-1995; 505509.

PR 30-JUL-1992; US-922987.

PR 31-AUG-1994; US-298087.

PA (UYNE-) UNIV NEW JERSEY.

PI Leibowitz MJ, Liu Y;

DR WPI; 98-398016/34.

PT Detection of Pneumocystis carinii - by amplification of nucleic acid

from sample with PCR primers specific for the 26S rRNA gene of

PT Pneumocystis carinii

PS Disclosure: Columns 29-34; 42pp; English.

CC The present sequence represents the total contiguous sequence of

CC Pneumocystis carinii in immunosuppressed Sprague-Dawley rats. The

CC specification describes a method for the diagnosis of Pneumocystis

CC carinii which comprises detecting the presence of a nucleic acid

CC sequence containing the 26S rRNA gene specific for P. carinii in a

CC sample. The 26S rRNA gene in a sample is amplified, and the primer

CC extension products detected by hybridisation with a labelled

CC oligonucleotide. The methods can be used for the diagnosis of

CC P. carinii infection and for the detection of various species of

CC P. carinii.

SQ Sequence 4256 BP; 1190 A; 766 C; 1124 G; 1176 T;

alignment_scores:

Quality: 29.00

Ratio: 4.833

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-5 x V36211/rev ..

Align seg 1/1 to reverse of: V36211 from: 1 to: 4256

1 ArgilleLeuLeuArgTyr 6

|||||

3792 CGAATCTGCTCGGTAT 3775

seq_name: N_Geneseq_36.V83003

seq_documentation_block:

ID V83003 standard; DNA; 4256 BP.

AC V83003;
 DT 23-FEB-1999 (first entry)
 DE Contiguous sequence determined for P. carinii from immunosuppressed rats.
 DE PCR amplification; sequencing; assay; inhibitor; nuclear rRNA gene;
 KW catalytic group I self-splicing intron reaction; drug screening; ss.
 OS Pneumocystis carinii.
 FH Key Location/Qualifiers
 FT Intron 1. .22
 FT /*tag= a
 FT /note= "3'-terminal portion of intron 1"
 FT 23. .53
 FT /*tag= b
 FT /note= "exon 2 of 16SrRNA"
 FT 54. .216
 FT /*tag= c
 FT /note= "internal transcribed spacer 1"
 FT 217. .374
 FT /*tag= d
 FT /note= "5.8S rRNA"
 FT 375. .556
 FT /*tag= e
 FT /note= "internal transcribed spacer 2"
 FT 557. .4256
 FT /*tag= f
 FT /note= "26S rRNA"

US5849484-A.

PD 15-DEC-1998.

PF 19-JUN-1995; 491690.

PR 27-MAY-1993; US-068248.

PR 30-JUL-1992; US-922987.

PR 19-JUN-1995; US-491690.

PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

PI Leibowitz MJ, Liu Y;

DR WPI; 99-069716/06.

PT Screening assays for drugs against Pneumocystis carinii - based on

inhibition of 26S rRNA gene intron self-splicing

PS Example 1; Fig 2; 51pp; English.

CC The present sequence represents the contiguous sequence determined for

CC Pneumocystis carinii from immunosuppressed Sprague-Dawley rats, using

CC the in vitro method of the invention. The method assays for an inhibitor

CC of the catalytic group I self-splicing intron reaction in the nuclear

CC rRNA genes of P. carinii. The method is useful for screening potential

CC drugs for treating P. carinii infections before more costly animal

CC testing is conducted.

SQ Sequence 4256 BP; 1190 A; 766 C; 1124 G; 1176 T;

alignment_scores:

Quality: 29.00

Ratio: 4.833

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-5 x V83003/rev ..

Align seg 1/1 to reverse of: V83003 from: 1 to: 4256

1 ArgilleLeuLeuArgTyr 6

|||||

3792 CGAATCTGCTCGGTAT 3775

seq_name: N_Geneseq_36.X12982

seq_documentation_block:

ID X12982 standard; DNA; 15614 BP.

AC X12982;

DT 19-MAR-1999 (first entry)

DE Enterococcus faecalis genome contig SEQ ID NO:45.

KW Enterococcus faecalis; contig; detection; Enterococcal infection;

KW vaccine; attenuation; computer readable medium; ds.

OS Enterococcus faecalis.

PN WO9850555-A2.

PD 12-NOV-1998.


```

PF 04-MAY-1998; U08985.
PR 14-NOV-1997; US-066009.
PR 06-MAY-1997; US-044031.
PR 16-MAY-1997; US-046655.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Dillon PJ, Kunsch CA;
DR WPI; 99-045171/04.
PT New isolated Enterococcus faecalis polynucleotides and polypeptides
PT - used to develop products for the detection of Enterococcus and for
PT use in vaccines for prevention or attenuation of Enterococcus
PT infection.
PS Claim 1; Page 419-427; 2084pp; English.
CC A computer readable medium has been developed which has recorded on it
CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
CC X12938 to X13919 represent these nucleotide sequences which are primary
CC nucleotide sequences, also known as contigs. The computer-based system
CC can identify fragments of the Enterococcus faecalis genome with
CC commercial importance. The products can be used to detect the presence
CC of Enterococcus faecalis in samples. They can also be used for
CC diagnosing Enterococcal infection in an animal and monitoring
CC progression of disease, and for identifying agents which can be used to
CC modulate the growth or pathogenicity of Enterococcus faecalis, or
CC another related organism, in vivo or in vitro. In particular the
CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
CC can be used in vaccines to prevent or attenuate an Enterococcal
CC infection.
SQ Sequence 15614 BP; 5535 A; 2567 C; 3169 G; 4326 T;

alignment_scores:
  Quality: 29.00      Length: 6
  Ratio: 4.833      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-5 x X12982/rev ..
Align seg 1/1 to reverse of: X12982 from: 1 to: 15614

1 ArgilleLeuLeuArgTyr 6
|||||
6777 CGAATATTATTCGATAC 6760

seq_name: N_Geneseq_36:Q26060
seq_documentation_block:
AC Q26060 standard; DNA; 51 BP.
DE 14-DEC-1992 (first entry)
DE SERP gene PCR primer p3.
KW Polymerase chain reaction; hybrid antigen; Plasmodium falciparum;
KW malaria; vaccine; ss.
OS Synthetic.
FH Key Location/Qualifier
FT misc_feature 1..18
FT /tag= a
FT /note= "complementary to SERP gene bases 3022-3045"
FT misc_feature 19..45
FT /tag= b
FT /note= "complementary to bases 226-252 of the
FT partial sequence 31-1 Ird of MSA I"
FT
FT
FT DE4041836-A.
PN 25-JUN-1992.
PF 24-DEC-1990; 041836.
PR 24-DEC-1990; DE-041836.
PA (BEHW ) BEHRINGERWKE AG.
PI Enders B, Hundt E, Knapp B, Kupper H;
DR WPI; 92-218121/27.
PT New plasmodium falciparum hybrid proteins for use in vaccines -
PT comprise at least two T-cell epitopes of SERP and a partial
PT sequence of HRPII
PS Example; Page 3; 10pp; German.
CC The sequence is that of a PCR primer used to obtain a 360bp fragment

```

```

CC which codes for amino acids 140-254 of the Plasmodium falciparum
CC antigen MSA I. The obtd. fragment is used in the construction of
CC hybrid antigen SERP/MSAI/HRPII which can be used in anti-malarial
CC vaccines which have a significantly stronger protective effect than
CC those using individual proteins. The vaccines are also easier to
CC purify and are less likely to be rendered ineffective by a mutation
CC of the pathogen. See Q26058-Q26061.
SQ Sequence 51 BP; 18 A; 6 C; 11 G; 16 T;

alignment_scores:
  Quality: 28.00      Length: 6
  Ratio: 4.667      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
US-08-653-294-5 x Q26060/rev ..
Align seg 1/1 to reverse of: Q26060 from: 1 to: 51

1 ArgilleLeuLeuArgTyr 6
|||||
31 AGAGTTCCTCTCGCTAT 14

seq_name: N_Geneseq_36:T26893
seq_documentation_block:
ID T26893 standard; DNA; 607 BP.
AC T26893;
DE 21-AUG-1996 (first entry)
DE Haemophilus influenzae 15 kD outer membrane protein coding sequence.
KW Outer membrane protein; immunoreactive; OMP; type b H.Influenzae;
KW non-typable H. influenzae; murine monoclonal antibody; vaccine; therapy;
KW passive immunisation; ss.
OS Haemophilus influenzae.
FH Key Location/Qualifiers
FT cds 171..542
FT /tag= a
FT /product= 15 kD outer membrane protein
FT
FT
FT US5503992-A.
PN 02-APR-1996.
PF 07-SEP-1993; 061314.
PR 07-SEP-1993; US-061314.
PA (CNDG ) CANADA MIN HEALTH.
PA (UNIW ) UNIV WASHINGTON.
PI Brodeur BR, Grass S, Hamel J, Munson RS;
DR WPI; 96-187695/19.
DR P-PSDB; R95448.
PT DNA coding for H. influenzae outer membrane protein - for prodn. of
PT recombinant polypeptide useful in vaccines, etc
PS Claim 1; Column 9-10; 11pp; English.
CC This sequence represents the coding sequence for a 15 kD outer membrane
CC protein (OMP) of Haemophilus influenzae. The 15 kD OMP is conserved
CC among type b and non-typable H. influenzae. Epitopes of the protein
CC sequence are recognised by murine monoclonal antibodies. The protein
CC encoded by this sequence can be used as a standard or reagent in
CC diagnostic tests. The protein can also be used in vaccines against type
CC b and non-typable H. influenzae. Monoclonal antibodies directed against
CC the encoded protein can be used as diagnostic reagents, and as
CC therapeutic agents for passive immunisation.
SQ Sequence 607 BP; 206 A; 97 C; 108 G; 196 T;

alignment_scores:
  Quality: 28.00      Length: 6
  Ratio: 4.667      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
US-08-653-294-5 x T26893/rev ..
Align seg 1/1 to reverse of: T26893 from: 1 to: 607

```

1 ArgileLeuLeuArgTyr 6
||||:|||||

64 AGGCTGCTACTAGATAC 47

seq_name: N_Geneseq_36:Q42561

seq_documentation_block:

ID Q42561 standard; DNA; 2960 BP.

AC Q42561;

DT 15-SEP-1993 (first entry)

DE Histamine H1 receptor coding sequence.

KW Bovine; histamine; H1; receptor; expression vector; pBH1; pEF-BOS-H1;

KW E. coli; allergy; ss.

OS Bos taurus.

FH Key Location/Qualifiers

FT cds 108.1583

FT /*tag= a

PN J05103674-A.

PD 27-APR-1993.

PF 17-OCT-1991; 269275.

PR 17-OCT-1991; JP-269275.

PA (OSAB-) ZH OSAKA BIOSCIENCE KENKYUSHO.

DR WPI; 93-171831/21.

DR P-PSDB; R36979.

PT New bovine histamine H1 receptor - used for treating allergy

PS Claim 2; Fig 1-5; lpp; Japanese.

CC This sequence encodes bovine histamine H1 receptor. This sequence

can be used in the construction of the expression vectors, pBH1 or

pEF-BOS-H1. These vectors may be used to transform E. coli, such

that they produce the H1 receptor protein. The histamine H1

receptor may be used for the treatment of allergy.

CC Sequence 2960 BP; 745 A; 743 C; 758 G; 714 T;

SQ

alignment_scores:

Quality: 28.00 Length: 6

Ratio: 4.667 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:

US-08-653-294-5 x Q42561/rev ..

Align seg 1/1 to reverse of: Q42561 from: 1 to: 2960

1 ArgileLeuLeuArgTyr 6

||||:|||||

538 CGGCTCTGTAGATAC 521

seq_name: N_Geneseq_36:T72653

seq_documentation_block:

ID T72653 standard; CDNA; 4617 BP.

AC T72653;

DT 20-AUG-1997 (first entry)

DE Rat tripeptidylpeptidase II cDNA clone.

KW Tripeptidylpeptidase II; Tpp II; rodent; rat; cerebral cortex;

KW neuropeptide; cholecystokinin; CCK; inactivation; degradation;

KW anorexia; schizophrenia; Parkinson's disease; depression;

KW irritable bowel syndrome; bulimia; pathological obesity;

KW alternative splicing; ss.

OS Rattus sp.

FH Key Location/Qualifiers

FT cds 65.3814

FT /*tag= a

FT /product= Tripeptidylpeptidase_II

FT misc_recomb 357.358

FT /*tag= b

FT /label= Alternative_splice_site

FT /note= "A second clone was isolated which differed

upstream of this site and was identical

downstream of it (see T72654)"

PN WO9635805-A2.

PD 14-NOV-1996.

PF 09-MAY-1996; F00700.
PR 09-MAY-1995; FR-005489.
PA (INRA) INSERM INST NAT SANTE & RECH MEDICALE.
PI Bantal RB, Bishop PB, Bourgeat P, Chan S, Ganellin CR;
PI Leblond B, Moore ANJ, Schwartz JC, Vargas F;
PI Lihua Z, Rose C;
DR WPI; 96-518693/51.
DR P-PSDB; W21571.
PT Screening medicaments for treating disorders linked to inactivation
PT of endogenous neuro-peptide(s) - by contacting candidate molecule
PT with membrane tri-peptidyl-peptidase (homologue) and measuring
PT enzyme activity
PS Disclosure: Pages 161-165; 212pp; French.
CC An enzyme with specificity for cholecystokinin (CCK) substrates
CC (specifically the non-sulphated CCK8 and the CCK5 peptides) was
CC purified from rat cerebral cortex membranes using high performance
CC liquid chromatography. Based on amino acid sequence data from the
CC purified enzyme, PCR primers were designed to amplify oligonucleotide
CC probes A and B of lengths 350 and 380 nucleotides, respectively.
CC Using these probes, two distinct clones were identified in a rat
CC brain cDNA library. The first clone had the present sequence and
CC is the rodent homologue of human tripeptidylpeptidase II. The
CC second clone (see T72653) differed in the 5' region, probably as a
CC result of alternative splicing, and appears to be a serine
CC ectopeptidase. The rat enzymes are preferred for use in a new
CC method of screening for medicaments for treating disorders linked
CC to the inactivation or degradation (or being treatable by retarding
CC physiological degradation) of endogenous neuropeptides. In the
CC method, a candidate molecule is contacted with a membrane tripeptidyl-
CC peptidase or homologue, and enzyme activity is measured.
CC Disorders linked to inactivation or degradation of endogenous
CC neuropeptides, include food intake disorders, cognitive and motor
CC disorders such as anorexia, schizophrenia, Parkinson's disease and
CC depression, as well as gastrointestinal transit disorders such as
CC irritable bowel syndrome, bulimia and pathological obesity.
SQ Sequence 4617 BP; 1429 A; 905 C; 999 G; 1284 T;

alignment_scores:

Quality: 28.00 Length: 6

Ratio: 4.667 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:

US-08-653-294-5 x T72653 ..

Align seg 1/1 to: T72653 from: 1 to: 4617

1 ArgileLeuLeuArgTyr 6

||||:|||||

2100 CGGCTCTGTAGATAT 2117

seq_name: N_Geneseq_36:T71315

seq_documentation_block:

ID T71315 standard; CDNA; 5072 BP.

AC T71315;

DT 10-SEP-1997 (first entry)

DE CDNA encoding C. elegans UNC-53 protein variant 7A.

KW UNC-53; neuronal regeneration; revascularisation; wound healing;

KW neurodegenerative disease; Alzheimer's disease; Huntington's;

KW peripheral neuropathies; metastasis inhibition; cancer; ss.

OS Caenorhabditis elegans.

FH Key Location/Qualifiers

FT cds 64.4815

FT /*tag= a

FT misc_feature 4594.4600

FT /*tag= b

FT /note= "region of seven A bases; the 8A UNC-53

variant has eight bases at this position

resulting in a different product"

PN WO9638555-A2.

PD 05-DEC-1996.

PF 31-MAY-1996; E02311.
 PR 31-MAY-1995; GB-010944.
 PA (BOGA/) BOGAERT T.
 PA (STRI/) STRINGHAM E.
 PA (VAND/) VANDEKERCKHOVE J.
 PI Bogaert T, Stringham E, Vandekerckhove J;
 DR WPI; 97-034369/03.
 DR P-PSDB; W20057.
 PT Caenorhabditis elegans UNC-53 protein 8A and 7A variants - useful to
 PT promote neuronal regeneration, revascularisation or wound healing.
 PS Claim 6; Page 103-106; 278pp; English.
 CC cDNA sequences encoding UNC-53 protein 8A and 7A variants of C. elegans
 CC are new. The UNC-53 proteins and nucleic acids are useful as medicaments
 CC to promote neuronal regeneration, revascularisation or wound healing, or
 CC for treatment of chronic neurodegenerative diseases (e.g. Alzheimer's or
 CC Huntington's disease) or acute traumatic injuries. Transgenic cells and
 CC organisms transfected with UNC-53 cDNA can be used to determine whether
 CC a substance is an inhibitor or enhancer of the regulation of cell shape
 CC or motility or the direction of cell migration by screening for a
 CC phenotypic change in the cell. Inhibitors can be used to alleviate the
 CC spread of disease inducing cells or metastasis. Probes derived from the
 CC cDNA sequences can be used to identify homologues of the C. elegans
 CC unc-53 gene. The UNC-53 protein can be used to identify proteins which
 CC are active in the signal transduction pathway that can be used as
 CC mentioned above.
 SQ Sequence 5072 BP; 1525 A; 1325 C; 1019 G; 1203 T;

alignment_scores:
 Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
 US-08-653-294-5 x T71315/rev ..

Align seg 1/1 to reverse of: T71315 from: 1 to: 5072

1 Arg1leLeuLeuArgTyr 6
 |||:::|||||
 1230 CGAGTCCTGTGAGATAT 1213

seq_name: N_Geneseq_36:T71314

seq_documentation_block:
 ID T71314 standard; cDNA; 5073 BP.
 AC T71314;
 DT 10-SEP-1997 (first entry)
 DE cDNA encoding C. elegans UNC-53 protein variant 8A.
 KW UNC-53; neuronal regeneration; revascularisation; wound healing;
 KW neurodegenerative disease; Alzheimer's disease; Huntington's;
 KW peripheral neuropathies; metastasis inhibition; cancer; ss.
 OS Caenorhabditis elegans.
 FT key Location/Qualifiers
 FT cds 64..4650
 FT /*tag= a
 FT /*tag= b
 FT /*tag= c
 FT /*tag= d
 FT /note= "region of eight A bases; the 7A UNC-53
 FT variant has only seven A bases at this
 FT position resulting in a different product"

PN W09638555-A2.
 PD 05-DEC-1996.
 PF 31-MAY-1996; E02311.
 PR 31-MAY-1995; GB-010944.
 PA (BOGA/) BOGAERT T.
 PA (STRI/) STRINGHAM E.
 PA (VAND/) VANDEKERCKHOVE J.
 PI Bogaert T, Stringham E, Vandekerckhove J;
 DR WPI; 97-034369/03.
 DR P-PSDB; W20056.
 PT Caenorhabditis elegans UNC-53 protein 8A and 7A variants - useful to
 PT promote neuronal regeneration, revascularisation or wound healing.

PS Claim 1; Page 101-103; 278pp; English.
 CC cDNA sequences encoding UNC-53 protein 8A and 7A variants of C. elegans
 CC are new. The UNC-53 proteins and nucleic acids are useful as medicaments
 CC to promote neuronal regeneration, revascularisation or wound healing, or
 CC for treatment of chronic neurodegenerative diseases (e.g. Alzheimer's or
 CC Huntington's disease) or acute traumatic injuries. Transgenic cells and
 CC organisms transfected with UNC-53 cDNA can be used to determine whether
 CC a substance is an inhibitor or enhancer of the regulation of cell shape
 CC or motility or the direction of cell migration by screening for a
 CC phenotypic change in the cell. Inhibitors can be used to alleviate the
 CC spread of disease inducing cells or metastasis. Probes derived from the
 CC cDNA sequences can be used to identify homologues of the C. elegans
 CC unc-53 gene. The UNC-53 protein can be used to identify proteins which
 CC are active in the signal transduction pathway that can be used as
 CC mentioned above.
 SQ Sequence 5073 BP; 1526 A; 1325 C; 1019 G; 1203 T;

alignment_scores:
 Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
 US-08-653-294-5 x T71314/rev ..

Align seg 1/1 to reverse of: T71314 from: 1 to: 5073

1 Arg1leLeuLeuArgTyr 6
 |||:::|||||
 1230 CGAGTCCTGTGAGATAT 1213

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:29:36 ; Search time 122.56 Seconds
(without alignments)
1.160 Million cell updates/sec

Title: US-08-653-294-6

Perfect score: 29

Sequence: 1 YRLIR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|----------|---------------------|
| 1 | 29 | 100.0 | 6 | 1 W47264 | Immunomodulatory p |
| 2 | 29 | 100.0 | 6 | 1 W33783 | Peptide #4 used in |
| 3 | 29 | 100.0 | 10 | 1 W47268 | Immunomodulatory p |
| 4 | 29 | 100.0 | 10 | 1 W47272 | Immunomodulatory p |
| 5 | 27 | 93.1 | 387 | 1 W89446 | A partial gIdAl pr |
| 6 | 27 | 93.1 | 389 | 1 R15428 | 3-acylation enzyme |
| 7 | 27 | 93.1 | 392 | 1 W37739 | Corn threonine dea |
| 8 | 27 | 93.1 | 621 | 1 W82842 | Helicobacter pylor |
| 9 | 27 | 93.1 | 625 | 1 W89445 | A gIdAl protein se |
| 10 | 27 | 93.1 | 724 | 1 R33081 | Bacillus thuringie |
| 11 | 27 | 93.1 | 764 | 1 W01897 | Nonsense-mediated |
| 12 | 27 | 93.1 | 1089 | 1 W01896 | Nonsense-mediated |
| 13 | 26 | 89.7 | 43 | 1 R58902 | Drosophila-12 cadh |
| 14 | 26 | 89.7 | 43 | 1 R37142 | Protocadherin clon |
| 15 | 26 | 89.7 | 783 | 1 W27113 | Rat spleen pro-hor |
| 16 | 25 | 86.2 | 174 | 1 Y06848 | H. felis ORF4 prot |
| 17 | 25 | 86.2 | 623 | 1 W89444 | Partial gIdAl prot |
| 18 | 25 | 86.2 | 637 | 1 W89443 | GIdAl protein sequ |
| 19 | 24 | 82.8 | 6 | 1 W47263 | Immunomodulatory p |
| 20 | 24 | 82.8 | 6 | 1 W33781 | Peptide #2 used in |
| 21 | 24 | 82.8 | 10 | 1 R61548 | Peptide fragment (|
| 22 | 24 | 82.8 | 10 | 1 W47266 | Immunomodulatory p |
| 23 | 24 | 82.8 | 10 | 1 W47270 | Immunomodulatory p |
| 24 | 24 | 82.8 | 12 | 1 R95429 | HLA-B*2702 84-79-84 |
| 25 | 24 | 82.8 | 12 | 1 W33798 | Peptide B2702.84-7 |
| 26 | 24 | 82.8 | 12 | 1 W33799 | Immunomodulating d |
| 27 | 24 | 82.8 | 20 | 1 R29909 | HLA-B*2702 CTL modu |
| 28 | 24 | 82.8 | 20 | 1 R92911 | HLA-B*2702 CTL modu |
| 29 | 24 | 82.8 | 20 | 1 R92907 | HLA-B*2702 CTL modu |
| 30 | 24 | 82.8 | 20 | 1 R95428 | HLA-B*2702 84-75-84 |
| 31 | 24 | 82.8 | 20 | 1 R95430 | HLA-B*2702 84-75/77 |
| 32 | 24 | 82.8 | 20 | 1 W33778 | Immunomodulating d |
| 33 | 24 | 82.8 | 20 | 1 W33779 | Immunomodulating d |
| 34 | 24 | 82.8 | 20 | 1 W33792 | Peptide B2702.84-7 |

35 24 82.8 69 1 W89901 Antigen 3 from clu
36 24 82.8 158 1 R63866 HPV18 E6/E7 protei
37 24 82.8 158 1 R79656 HPV-18 E6 protein.
38 24 82.8 205 1 R39333 EpiQ protein. Nove
39 24 82.8 225 1 W89889 Antigen 1 from clu
40 24 82.8 225 1 W89840 Protein encoded by
41 24 82.8 271 1 R27728 HPV 18 E6 protein
42 24 82.8 278 1 Y02641 Prot.D1/3-E6-His/H
43 24 82.8 383 1 W41592 Rat FRAG1 protein.
44 24 82.8 383 1 Y02642 Prot.D1/3-E6-E7-H1
45 24 82.8 434 1 R95698 Erythrobacter long

ALIGNMENTS

RESULT 1
W47264
ID W47264 standard; peptide; 6 AA.
AC W47264;
DT 22-MAY-1998 (first entry)
DE Immunomodulatory peptide.
KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
KW transplant rejection; treatment; autoimmune disease.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 1..6
FT D-isomer
FT W09744052-A1.
PD 27-NOV-1997.
PF 23-APR-1997: U06705.
PR 22-MAY-1996; US-651650.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM;
DR WPI; 98-018220/02.
PT Novel immunomodulatory peptide-type compound - useful for inhibiting
PT transplant rejection
PS Claim 10; page 36; 41pp; English.
CC The present sequence is an immunomodulatory peptide, which
CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
CC in a pharmaceutical composition together with a subtherapeutic dose
CC of an immunosuppressant, to extend the period of acceptance of a
CC transplant from a major histocompatibility complex (MHC) unmatched
CC donor, i.e. to inhibit transplant rejection. It can also be used in
CC the treatment of autoimmune diseases.
CC Peptides using the D-form amino acids are more effective
CC immunomodulators than their diastereomers or enantiomers.
SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YRLIR 6
Db 1 YRLIR 6

RESULT 2
W33783
ID W33783 standard; peptide; 6 AA.
AC W33783;
DT 19-JUN-1998 (first entry)
DE Peptide #4 used in immunomodulating dimer peptide.
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
KW transplant rejection; autoimmune disease; Class I HLA-B alpha-1 domain;
KW rejection.
OS Synthetic.
OS Homo sapiens.
PN W09744351-A1.

PD 27-NOV-1997.
 PF 22-MAY-1997; U08689.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI: 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Claim 15; Page 35; 41pp; English.
 CC This sequence represents a specifically claimed peptide which forms part
 CC of the immunomodulating dimer peptides of the invention. A peptide-type
 CC compound or variant is claimed which has immunomodulating activity,
 CC including the N-terminal acylated and/or C-terminal amidated or
 CC esterified forms of up to 60 amino acids, where the peptide-type compound
 CC comprises the formula: A-B, where A, B = (R aa76-77L) (aa79-84) or
 CC (aa84-79) (Laa77-76R); aa76 - E or V; aa77 - D, S or N; aa79 - R or G;
 CC aa80 - I or N; aa81, aa84 - a hydrophobic or small amino acid; aa82 - R
 CC or L; aa83 - G or R; and aa represents amino acid. The sequence in the
 CC brackets may optionally be absent or truncated at any peptide type bond
 CC within the brackets. The compounds comprise amino acid sequences related
 CC to a Class I HLA-B alpha domain (positions 79-84). They can be used to
 CC inhibit cytotoxic T-lymphocytes (CTL) from undesirably attacking cells in
 CC a host or in vitro. They can also be used in combination with antigenic
 CC peptides or proteins of interest to activate CTLs. They can also inhibit
 CC the proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.
 SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIR 6
 |||||
 Db 1 YRLIR 6

RESULT 3

W47268
 ID W47268 standard; peptide; 10 AA.
 AC W47268;
 DT 22-MAY-1998 (first entry)
 DE Immunomodulatory peptide.
 KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
 KW transplant rejection; treatment; autoimmune disease.
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc_difference 1..10 /note- "at least one of the amino acids is the
 FT D-isomer

Query Match 100.0%; Score 29; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIR 6
 |||||
 Db 1 YRLIR 6

RESULT 5

W89446
 ID W89446 standard; Protein; 387 AA.
 AC W89446;
 DT 25-MAY-1999 (first entry)
 DE A partial gldal protein sequence.
 KW Gldal; bacterial infection; meningitis; Helicobacter pylori infection;
 KW cancer; ulcer; gastritis; antibacterial; in-dwelling device;
 KW wound treatment; bacterial adhesion; matrix protein.
 OS Staphylococcus aureus.
 FH Key Location/Qualifiers
 FT Misc_difference 283 /note- "unspecified amino acid encoded by ATN"
 FT EP-889129-A2.
 PD 07-JAN-1999.

CC Immunomodulators than their diastereomers or enantiomers.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.52;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIR 6
 |||||
 Db 1 YRLIR 6

RESULT 4

W47272
 ID W47272 standard; peptide; 10 AA.
 AC W47272;
 DT 22-MAY-1998 (first entry)
 DE Immunomodulatory peptide.
 KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
 KW transplant rejection; treatment; autoimmune disease.
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc_difference 1..10 /note- "at least one of the amino acids is the
 FT D-isomer

QY 1 YRLIR 6
 |||||
 Db 1 YRLIR 6

RESULT 5

W89446
 ID W89446 standard; Protein; 387 AA.
 AC W89446;
 DT 25-MAY-1999 (first entry)
 DE A partial gldal protein sequence.
 KW Gldal; bacterial infection; meningitis; Helicobacter pylori infection;
 KW cancer; ulcer; gastritis; antibacterial; in-dwelling device;
 KW wound treatment; bacterial adhesion; matrix protein.
 OS Staphylococcus aureus.
 FH Key Location/Qualifiers
 FT Misc_difference 283 /note- "unspecified amino acid encoded by ATN"
 FT EP-889129-A2.
 PD 07-JAN-1999.

Query Match 100.0%; Score 29; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.52;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIR 6
 |||||
 Db 1 YRLIR 6

RESULT 5

W89446
 ID W89446 standard; Protein; 387 AA.
 AC W89446;
 DT 25-MAY-1999 (first entry)
 DE A partial gldal protein sequence.
 KW Gldal; bacterial infection; meningitis; Helicobacter pylori infection;
 KW cancer; ulcer; gastritis; antibacterial; in-dwelling device;
 KW wound treatment; bacterial adhesion; matrix protein.
 OS Staphylococcus aureus.
 FH Key Location/Qualifiers
 FT Misc_difference 283 /note- "unspecified amino acid encoded by ATN"
 FT EP-889129-A2.
 PD 07-JAN-1999.

PF 30-JUN-1998; 305180.
 PR 01-JUL-1997; US-052758.
 PA (SWIK) SMITHKLINE BEECHAM CORP.
 PI (SWIK) SMITHKLINE BEECHAM PLC.
 PI Burnham M, Kallender H, Lenox AL, Palmer LM;
 DR N-PSDB; V82078.
 PT New isolated gldA polypeptide from *Staphylococcus aureus* - used to
 PT diagnose, treat and prevent bacterial infections e.g. *S. aureus* and
 PT *H. pylori* and associated cancers, ulcers and gastritis
 PS Claim 1; Page 6; 43pp; English.
 CC The present sequence represents a partial gldA protein of *Staphylococcus*
 CC *aureus*. GldA proteins, nucleic acids and agonists are used to
 CC treat conditions requiring increased activity or expression of gldA,
 CC while conditions (particularly bacterial infections) requiring
 CC inhibition of gldA are treated by administering an antagonist,
 CC inhibitory nucleic acid or competitive polypeptide. The products are
 CC used to treat *S. pneumoniae* infection, particularly meningitis and
 CC also *Helicobacter pylori* infections e.g. related cancers, ulcers and
 CC gastritis. These antibacterial agents may also be used to treat
 CC in-dwelling devices to prevent infection or generally as wound
 CC treatments to prevent adhesion of bacteria to matrix proteins.
 SQ Sequence 387 AA;

Query Match 93.1%; Score 27; DB 1; Length 387;
 Best Local Similarity 83.3%; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6
 |||||
 Db 196 YRLLR 201

RESULT 6
 RI5428
 ID RI5428 standard; Protein; 389 AA.
 AC RI5428;
 DE 25-FEB-1992 (first entry)
 DE 3-acylation enzyme.
 KW Macrolide antibiotics; acylation; Streptomycetes; tylosin;
 KW spiramycin; leucomycin; nidamycin; acyA.
 OS Streptomycetes thermotolerans ATCC 11416.
 PN EP-459525-A.
 PD 04-DEC-1991.
 PF 03-JUN-1991; 109033.
 PR 01-JUN-1990; JP-141625.
 PR 22-FEB-1991; JP-048753.
 PA (SAOC) MERCIAN CORP.
 PA Arisawa A, Kawamura N, Kojima I, Tone H, Okamoto R;
 PI Okamura K;
 DR WPI; 91-355809/49.
 DR N-PSDB; Q14978.
 PT Gene encoding 3-acylation enzyme for macrolide antibiotics - used
 PT to transform Streptomycetes sp. to have acylation activity, for use
 PT in medicine.
 PS Disclosure; Fig 7; 45pp; English.
 CC The sequence was deduced from DNA sequence obtd. from plasmid p53A,
 CC a subclone of a DNA library prepd. from genomic DNA of *S. thermotol-*
 CC *erans*. The 3-acylation enzyme is involved in the synthesis of
 CC macrolide antibiotics. Microorganisms, ep. Streptomycetes, trans-
 CC formed with the gene can be used for the commercial prodn. of
 CC macrolide antibiotics, e.g. 3-acylated tylosin.
 SQ Sequence 389 AA;

Query Match 93.1%; Score 27; DB 1; Length 389;
 Best Local Similarity 83.3%; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6
 |||||
 Db 158 YRLLR 163

RESULT 7
 W97739
 ID W97739 standard; Protein; 392 AA.
 AC W97739;
 DE 21-MAY-1999 (first entry)
 DE Corn threonine deaminase.
 KW Threonine deaminase; corn; maize; amino acid; lysine; threonine;
 KW methionine; cysteine; isoleucine; transgenic plant;
 KW crop improvement; food; feedstuff.
 OS Zea mays.
 PN WO9855601-A2.
 PD 10-DEC-1998.
 PF 05-JUN-1998; U11692.
 PR 12-JUN-1997; US-049443.
 PR 08-JUN-1997; US-048771.
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 PI Abell LM, Allen SM, Falco SC, Hitz WD, Kinney AJ,
 PI Rafalski JA, Thorpe CJ;
 DR WPI; 99-070263/06.
 DR N-PSDB; X07180.
 PT New plant amino acid biosynthetic enzymes, DNA and chimeric genes -
 PT encode: dihydropicolinate reductase; diaminopimelate epimerase;
 PT threonine synthase; threonine deaminase; S-adenosylmethionine
 PT synthetase
 PS Claim 31; Page 57-58; 98pp; English.
 CC This is the amino acid sequence of a near full-length corn
 CC threonine deaminase, as deduced from a cDNA clone (see X07180),
 CC designated cen1.pk0064.f4, obtained from a corn endosperm cDNA
 CC library. The amino acid sequence shows similarity to the
 CC Burkholderia capacia enzyme. The invention relates to new isolated
 CC nucleic acid fragments (see X07168-85) encoding plant enzymes (see
 CC W97727-44) that catalyze steps in the biosynthesis of lysine, the
 CC threonine, methionine, cysteine and isoleucine from aspartate, the
 CC enzyme being selected from dihydropicolinate reductase,
 CC diaminopimelate epimerase, threonine synthase, threonine deaminase
 CC or S-adenosylmethionine synthetase. The invention also relates to
 CC the construction of a chimeric gene encoding all or a portion of
 CC the biosynthetic pathway enzyme, in sense or antisense orientation,
 CC where expression of the chimeric gene results in production of
 CC altered levels of the enzyme in a transformed host cell.
 CC Overexpression or reduction of expression of genes encoding the
 CC amino acid biosynthetic pathway enzymes in crop plants such as
 CC corn, soybean and wheat can be used to alter levels of the amino
 CC acids in human food and animal feed. Transformed host cells can
 CC also be used to identify compounds that inhibit one of the enzymes.
 SQ Sequence 392 AA;

Query Match 93.1%; Score 27; DB 1; Length 392;
 Best Local Similarity 83.3%; Pred. No. 63;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6
 |||||
 Db 383 YRLLR 388

RESULT 8
 W62842
 ID W62842 standard; Protein; 621 AA.
 AC W62842;
 DE 26-OCT-1998 (first entry)
 DE *Helicobacter pylori* DapE protein.
 KW N-succinyl-L-diaminopimelic acid desuccinylase; DapE; vaccine;
 KW immunisation.
 OS *Helicobacter pylori* strain 60190.
 PN WO9827819-A1.
 PD 02-JUL-1998.
 PF 23-DEC-1997; U24147.
 PR 23-DEC-1996; US-033824.
 PA (UYVA-) UNIV VANDERBILT.

PI Blaser MJ, Karita M;
 DR WPI: 98-377287/32;
 DR N-PSDB; V42319 and V42322.
 PT Helicobacter pylori dape gene and related protein - used to create a
 PT mutant used in immunisation against infection by e.g. HIV,
 PT influenza, respiratory syncytial virus etc
 PS Example: Page 45-47; 77pp; English.
 CC This N-succinyl-L-diaminopimelic acid desuccinylase is encoded by
 CC the newly isolated dape gene (see V42319 and V42322) of
 CC Helicobacter pylori strain 60190. It shows 37.9% identity and
 CC 61.0% similarity with the dape protein of E. coli. Conditionally
 CC lethal dape- mutant H. pylori strains are provided; a claimed dape-
 CC mutant is deposited as ATCC 55997. Such mutants can be used in a
 CC claimed method of immunising a subject against H. pylori infection.
 CC The dape- mutants can also be used as hosts of nucleic acids
 CC encoding foreign proteins. Such mutants can be used in claimed
 CC methods of immunising against bacterial, viral, fungal or parasite
 CC infections. Sperm antigens may also be expressed for use in birth
 CC control.
 SQ Sequence 621 AA;

Query Match 93.1%; Score 27; DB 1; Length 621;
 Best Local Similarity 83.3%; Pred. No. 1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6

|||||

Db 432 YRLIIR 437

RESULT 9

W89445
 ID W89445 standard; Protein; 625 AA.

AC W89445;
 DE A GidAl protein sequence.
 KW GidAl; bacterial infection; meningitis; Helicobacter pylori infection;
 KW cancer; ulcer; gastritis; antibacterial; in-dwelling device;
 KW wound treatment; bacterial adhesion; matrix protein.
 OS Staphylococcus aureus.
 PN EP-889129-A2.
 PD 07-JAN-1999.
 PF 30-JUN-1998; 305180.
 PR 01-JUL-1997; US-052758.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 PI Burnham M, Kallender H, Lenox AL, Palmer LM;
 DR WPI: 99-062660/06.
 DR N-PSDB: V82077.
 PT New isolated gidAl polypeptide from Staphylococcus aureus - used to
 PT diagnose, treat and prevent bacterial infections e.g. S. aureus and
 PT H. pylori and associated cancers, ulcers and gastritis
 PS Claim 1: Page 4-5; 43pp; English.
 CC The present sequence represents a gidAl protein of Staphylococcus
 CC aureus. GidAl proteins, nucleic acids and agonists are used to
 CC treat conditions requiring increased activity or expression of gidAl,
 CC while conditions (particularly bacterial infections) requiring
 CC inhibition of gidAl are treated by administering an antagonist,
 CC inhibitory nucleic acid or competitive polypeptide. The products are
 CC used to treat S. pneumoniae infection, particularly meningitis and
 CC also Helicobacter pylori infections e.g. related cancers, ulcers and
 CC gastritis. These antibacterial agents may also be used to treat
 CC in-dwelling devices to prevent infection or generally as wound
 CC treatments to prevent adhesion of bacteria to matrix proteins.
 SQ Sequence 625 AA;

Query Match 93.1%; Score 27; DB 1; Length 625;
 Best Local Similarity 83.3%; Pred. No. 1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6

Db 432 YRLIIR 437

|||||

RESULT 10

R93081
 ID R93081 standard; Protein; 724 AA.
 AC R93081;
 DE Bacillus thuringiensis insecticidal protein JEG80.
 KW Insecticidal crystal toxin; JEG80; anti-diptera; mosquito; larvae;
 KW Aedes aegypti; Anopheles stephensi; Culex pipiens; BtJeg 367.
 OS Bacillus thuringiensis ser. JEG80.
 PN WO9606171-A2.
 PD 29-FEB-1996.
 PF 24-AUG-1995; FO1116.
 PR 25-AUG-1994; FR-010299.
 PA (INSP) INST PASTEUR.
 PI Delecluse A, Thliery I;
 DR WPI: 96-151374/15.
 DR N-PSDB: T17044.
 PT New B.thuringiensis ser. JEG80 protein which was isolated
 PT related DNA, with high activity against dipteran larvae, e.g.
 PT A.aegypti, A.stephensi or C.pipiens
 PS Claim 17; Fig 5; 55pp; French.
 CC The present sequence is that of the JEG80 protein which was isolated
 CC from Bacillus thuringiensis ser. JEG80 (BtJeg367); it is
 CC a crystal toxin of mol. wt. 80 kD which is active against dipteran
 CC insects, esp. mosquitoes. The full-length, recombinant JEG80 toxin,
 CC had mean LC50 values (in ng/ml) after 48 hours of 18.8, 42.7 and
 CC 10.1 against larvae of Aedes aegypti, Anopheles stephensi and Culex
 CC pipiens, respectively. Wild-type BtJeg367 crystals gave values of
 CC 47.4, 54.5 and 9.6. The JEG80 protein is far more toxic (6-40 times
 CC more toxic against the mosquito species tested) than the Bt CryIIVD
 CC toxin, despite their close sequence similarity.
 SQ Sequence 724 AA;

Query Match 93.1%; Score 27; DB 1; Length 724;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6

|||||

Db 548 YRLIIR 553

RESULT 11

W01897
 ID W01897 standard; Protein; 764 AA.
 AC W01897;
 DE 24-NOV-1996 (first entry)
 DE Nonsense-mediated mRNA decay 2 C-terminal.
 DE Nonsense-mediated mRNA decay; NMD2; Upflp; inhibition.
 OS Saccharomyces cerevisiae.
 PN WO9622301-A1.
 PD 25-JUL-1996.
 PF 27-DEC-1995; U16930.
 PR 20-JAN-1995; US-375300.
 PA (UYMA-) UNIV MASSACHUSETTS MEDICAL CENT.
 PI He F, Jacobson AS;
 DR WPI: 96-354469/35.
 DR N-PSDB: T31994.
 PT Isolated nonsense-mediated mRNA decay gene and protein - used to
 PT develop methods for inhibiting the decay pathway for producing
 PT heterologous or endogenous proteins
 PS Claim 10; Page 54-56; 73pp; English.
 CC The NMD2 gene (T31993) is named after its role in the Nonsense-
 CC mediated mRNA Decay pathway. The protein, Nmd2p, binds to Upflp.
 CC A C-terminal fragment of the protein (T31994) also binds Upflp and,
 CC when overexpressed in the host cell, the fragment inhibits the
 CC function of Upflp, thereby inhibiting the nonsense-mediated
 CC mRNA decay pathway.

SQ Sequence 764 AA;

Query Match 93.1%; Score 27; DB 1; Length 764;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIR 6
 DB 255 YRLIR 260
 ||:||||

RESULT 12

W01896
 ID W01896 standard; Protein; 1089 AA.
 AC W01896;
 DT 24-NOV-1996 (first entry)
 DE Nonsense-mediated mRNA decay 2 protein.
 KW Nonsense-mediated mRNA decay; NMD2; Upflp; inhibition.
 OS Saccharomyces cerevisiae.
 PN W09622301-AL.
 PD 25-JUL-1996.
 PF 27-DEC-1995; US-375300.
 PR 20-JAN-1995; US-375300.
 PA (UYMA-) UNIV MASSACHUSETTS MEDICAL CENT.
 PI He F, Jacobson AS;
 DR WPI; 96-354489/35.
 DR N-PSDB; T31993.
 PT Isolated nonsense-mediated mRNA decay gene and protein - used to develop methods for inhibiting the decay pathway for producing heterologous or endogenous proteins
 PS Claim 8: Page 49-52; 73pp; English.
 CC The NMD2 gene (T31993) is named after its role in the Nonsense-mediated mRNA decay pathway. The protein, Nmd2p, binds to Upflp.
 CC A C-terminal fragment of the protein (T31994) also binds Upflp.
 CC When overexpressed in the host cell, the fragment inhibits the function of Upflp, thereby inhibiting the nonsense-mediated mRNA decay pathway.
 CC Sequence 1089 AA;

Query Match 93.1%; Score 27; DB 1; Length 1089;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIR 6
 DB 580 YRLIR 595
 ||:||||

RESULT 13

R58902
 ID R58902 standard; Protein; 43 AA.
 AC R58902;
 DT 17-APR-1995 (first entry)
 DE Drosophila-12 cadherin-related molecule.
 KW Cadherin; cell adhesion molecule.
 OS Drosophila.
 PN W09414960-A.
 PD 07-JUL-1994.
 PF 23-DEC-1993; U12588.
 PR 29-DEC-1992; US-998003.
 PA (DOHE-) DOHENY EYE INST.
 PI Suzuki S;
 DR WPI; 94-293849/36.
 DR N-PSDB; Q68993.
 PT Polynucleotide sequences encoding new proto:cadherins - useful for modulating natural binding and regulating activities.
 PS Example; Page 63; 114pp; English.
 CC Two regions of conserved AA sequence, one from the middle of the third cadherin extracellular subdomain (EC-3) and the other from the C-terminus of the fourth extracellular subdomain (EC-4) were identified. The corresp. degenerate oligos (Q68949, Q68950) were

CC designed for use as PCR primers. PCR was carried out on a rat brain cDNA prep. Two major bands of about 450 bps and 130 bps were found. The 450 bp band corresponded to the expected length between the two primer sites, but the 130 bp band could not be predicted from any of the previously identified cadherin sequences. The 450 bp and 130 bp bands were extracted and sequenced. Nineteen novel partial cDNA clones were isolated. The DNA and deduced AA sequences of the clones (including sequences corresp. to the PCR primers) are given in Q68951-Q68969 and R58860-R58878. Various cDNA fragments structurally similar to the rat cDNAs were isolated from human, mouse and xenopus brain cDNA preps. and from Drosophila and C. elegans whole body cDNA preps. by PCR using the above primers. The DNA and deduced AA sequences of the resulting PCR fragments (including sequences corresp. to the PCR primers) are given in Q68971, Q68972-Q68994 and R58882-R58905 and R49143. Comparison of the deduced AA sequences indicates a similarity, in particular, there are three sets of clones that appear to be cross-species homologues: RAT-218, MOUSE-322 and HUMAN-43; RAT-314, MOUSE-321 and HUMAN-11; and MOUSE-325 and HUMAN-42.
 SQ Sequence 43 AA;

Query Match 89.7%; Score 26; DB 1; Length 43;
 Best Local Similarity 83.3%; Pred. No. 11;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIR 6
 DB 11 YRLIR 16
 |||:||||

RESULT 14

R87142
 ID R87142 standard; Peptide; 43 AA.
 AC R87142;
 DT 29-AUG-1996 (first entry)
 DE Protocadherin clone DROSOPHILA-12.
 KW Protocadherin; pc3; pc4; pc5; human; rat; cadherin; cell adhesion; mouse; catenin; therapy; clone; frog; fruit fly.
 OS Drosophila melanogaster.
 FH Key Location/Qualifiers
 FT misc_difference 38 /note= "encoded by CGA"
 FT misc_difference 39 /note= "encoded by ATG"
 FT misc_difference 40 /note= "encoded by ACA"
 FT misc_difference 41 /note= "encoded by ATG"
 FT misc_difference 42 /note= "encoded by CGC"
 FT misc_difference 43 /note= "encoded by CGC"
 FT W09600289-AL.
 PN 04-JAN-1996.
 PD 26-JUN-1995; U08071.
 PR 27-JUN-1994; US-268161.
 PA (DOHE-) DOHENY EYE INST.
 PI Suzuki S;
 DR WPI; 96-068873/07.
 DR N-PSDB; T03617.
 PT Polynucleotide(s) encoding human protocadherins pc3 and pc4 and rat pc5 - involved in cell-cell adhesion and regulation activities
 PS Example 2; Page 57-68; 146pp; English.
 CC R87142-R87144 represent partial fragments of the drosophila protocadherin sequence. The cDNAs encoding these sequences were isolated after screening a drosophila whole body cDNA preparation with the primers shown in T03575 and T03576. The primers were constructed from portions of the amino acid sequences of the third and fourth extracellular domains of published cadherin sequences. The cytoplasmic domain of cadherin interacts with the cytoskeleton through catenins and other cytoskeleton associated proteins. The cytoplasmic domain is not present in all cadherins, but in those which possess it, it is essential for the

CC cadherins adhesive function. The cadherins which do not possess a
 CC cytoplasmic domain appear to function via a different method from those
 CC with a cytoplasmic domain. These protein sequences are involved in
 CC cell-cell adhesion. These sequences may have regulatory functions in the
 CC cell, as well as the cell-cell adhesive properties. Antibodies produced
 CC against these sequences are useful for modulating the binding activity of
 CC these protocadherins, and can be used therapeutically.
 SQ Sequence 43 AA;

Query Match 89.7%; Score 26; DB 1; Length 43;

Best Local Similarity 83.3%; Pred. No. 11;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6

|||:|

Db 11 YRLVIR 16

RESULT 15

W27113
 ID W27113 standard; Protein: 783 AA.

AC W27113;

DE 20-NOV-1997 (first entry)

KW Rat spleen pro-hormone convertase 7.

KW PC7; HIV gp160; human immunodeficiency virus; Alzheimer's disease;

KW Down's syndrome; AIDS; inhibitor; antibody; subtilisin-kexin protein;

OS lymphoid tissue; CD4+ cell; rat; human.

OS Rattus rattus.

FH Key

FT peptide

FT 1..36

FT /label= Signal

FT 37..783

FT /label= PC7

FT 137..140

FT /note= "Putative zymogen activation site"

FT modified_site

FT 166

FT /note= "Putative glycosylation site"

FT modified_site

FT 174

FT /note= "Putative glycosylation site"

FT active_site

FT 186

FT active_site

FT 227

FT modified_site

FT 240

FT /note= "Putative glycosylation site"

FT misc_difference

FT 328

FT /label= Oxyanion_hole

FT active_site

FT 405

FT modified_site

FT 510

FT /note= "Putative glycosylation site"

FT modified_site

FT 540

FT /note= "Putative phosphorylation"

FT domain

FT 667..683

FT /note= "Putative transmembrane anchoring domain"

FT WO9705256-A2.

PN 13-FEB-1997.

PD 02-AUG-1996; CA0520.

PF 19-OCT-1995; US-545562.

PR 02-AUG-1995; US-510347.

PR 18-AUG-1995; US-517015.

PA (RECL-) INST RECH CLINIQUES MONTREAL.

PI Chretien M, Day R, Seidah NG;

DR WPI: 97-145693/13.

DR N-PSDB; T85113.

PT Human and rat pro-hormone convertase, PC7, can cleave HIV gp160 -

PT may allow early diagnosis of Alzheimer's disease and Down's

PT syndrome, also PC7 inhibitors are used to prevent development of

PT AIDS

PS Claim 1: Page 29-31; 52pp; English.

CC The present sequence represents a purified rat pro-hormone convertase,

CC designated PC7. PC7 is a member of the subtilisin-kexin protein family,

CC it is widely distributed (especially in lymphoid tissue) and is a

CC possible maturation enzyme for gp160 in target CD4+ cells. Since PC7

CC has been found in many brains of patients with Alzheimer's disease and

CC Down's syndrome, it may allow early diagnosis of these diseases. It is
 CC also contemplated that over-expression of PC7 may protect against the
 CC disease. Antibodies against PC7 are used to detect/quantify PC7 in
 CC cells and tissues, and oligonucleotides able to hybridize to a PC7
 CC coding sequence, are used similarly at the nucleic acid (especially
 CC mRNA) level. Oligonucleotides can also be used to amplify PC7-encoding
 CC nucleic acid. Recombinant host cells are used to produce PC7 which is
 CC able to convert precursors to mature proteins, especially HIV gp160 to
 CC g120 and gp41. Antisense oligonucleotides, or vectors expressing them,
 CC are used to prevent expression of PC7 in cells susceptible to HIV
 CC infection; they prevent production of gp120 and so development of HIV
 CC infection to AIDS.
 SQ Sequence 783 AA;

Query Match 89.7%; Score 26; DB 1; Length 783;

Best Local Similarity 83.3%; Pred. No. 2.2e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6

|||:|

Db 587 YRLVIR 592

Search completed: February 8, 2000, 01:29:36

Job time: 1748 sec

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 11:54:16 ; Search time 117.7 seconds
(without alignments)
2.405 Million cell updates/sec

Title: US-08-653-294-6

Perfect score: 29

Sequence: 1 YRLIIR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

PIR62.*

1: pirl.*

2: pirl2.*

3: pirl3.*

4: pirl4.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result NO. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 28 | 96.6 | 325 | 2 B72475 | probable transcrip |
| 2 | 28 | 96.6 | 1489 | 2 S60416 | DNA helicase YGL15 |
| 3 | 27 | 93.1 | 101 | 2 S50445 | hypothetical prote |
| 4 | 27 | 93.1 | 157 | 2 H75034 | hypothetical prote |
| 5 | 27 | 93.1 | 221 | 2 S56263 | hypothetical prote |
| 6 | 27 | 93.1 | 239 | 2 S71158 | hypothetical prote |
| 7 | 27 | 93.1 | 389 | 2 JC4001 | macrolide 3-O-acyl |
| 8 | 27 | 93.1 | 621 | 2 D71961 | glucose inhibited |
| 9 | 27 | 93.1 | 621 | 2 E64546 | glucose-inhibited |
| 10 | 27 | 93.1 | 628 | 1 BWBSGA | glucose-inhibited |
| 11 | 27 | 93.1 | 629 | 2 I64078 | gida protein - Hae |
| 12 | 27 | 93.1 | 629 | 2 F72400 | glucose-inhibited |
| 13 | 27 | 93.1 | 720 | 1 A48723 | replication licens |
| 14 | 27 | 93.1 | 1021 | 2 T10748 | mannan endo-1,4-be |
| 15 | 27 | 93.1 | 1089 | 2 S48244 | NMD2 protein - yea |
| 16 | 26 | 89.7 | 116 | 2 F72735 | hypothetical prote |
| 17 | 26 | 89.7 | 173 | 2 S27599 | hypothetical prote |
| 18 | 26 | 89.7 | 233 | 2 B69202 | endonuclease III - |
| 19 | 26 | 89.7 | 357 | 2 A71869 | integrase/recombin |
| 20 | 26 | 89.7 | 374 | 2 A71320 | hypothetical prote |
| 21 | 26 | 89.7 | 447 | 2 T16527 | hypothetical prote |
| 22 | 26 | 89.7 | 469 | 2 F71307 | hypothetical prote |
| 23 | 26 | 89.7 | 783 | 2 JC6136 | kexin-like protein |
| 24 | 26 | 89.7 | 785 | 2 S64706 | subtilisin-like pr |
| 25 | 25 | 86.2 | 145 | 2 T13342 | hypothetical prote |
| 26 | 25 | 86.2 | 234 | 2 JH0421 | proteasome chain x |
| 27 | 25 | 86.2 | 234 | 2 S36116 | multicatalytic end |
| 28 | 25 | 86.2 | 271 | 2 T08009 | probable ribosomal |
| 29 | 25 | 86.2 | 297 | 2 S42144 | ribosomal protein |
| 30 | 25 | 86.2 | 297 | 2 T12615 | ribosomal protein |

31 25 86.2 380 2 D72684
32 25 86.2 416 2 A56486
33 25 86.2 605 2 H71303
34 25 86.2 629 1 BVECOA
35 25 86.2 630 1 BWPSAP
36 25 86.2 635 2 S76371
37 25 86.2 1075 2 T07448
38 24 82.8 81 2 S47689
39 24 82.8 87 2 D70886
40 24 82.8 96 1 IECE99
41 24 82.8 149 2 E71728
42 24 82.8 153 2 A36837
43 24 82.8 156 1 D71689
44 24 82.8 158 1 W6WL18
45 24 82.8 158 1 W6WL39

ALIGNMENTS

RESULT 1
B72475
Probable transcription initiation factor IIB APE2443 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: B72475
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339
A:Accession: B72475
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <KAW>
A:Cross-references: DDBJ:AP000064; NID:g5105945; PIDN:BA81458.1; PID:d1045244; PID:g
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE2443

Query Match 96.6%; Score 28; DB 2; Length 325;
Best Local Similarity 83.3%; Pred. No. 21;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6
| | | | |
Db 209 YRLIIR 214

RESULT 2
S60416
DNA helicase YGL150c - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein G1880
C:Species: Saccharomyces cerevisiae
C:Date: 15-Feb-1996 #sequence_revision 13-Apr-1996 #text_change 21-Nov-1997
C:Accession: S60416; S53049; S60421; S64166; S64164
R:James, C.M.; Indge, K.J.; Oliver, S.G.
submitted to the EMBL Data Library, March 1995
A:Description: DNA sequence analysis of a 35kb segment from Saccharomyces cerevisiae
A:Reference number: S60416
A:Accession: S60416
A:Molecule type: DNA
A:Residues: 1-1489 <JAM>
A:Cross-references: EMBL:Z48618
R:James, C.M.; Indge, K.J.; Oliver, S.G.
submitted to the EMBL Data Library, March 1995
A:Description: DNA sequence analysis of a 35kb segment from S.cerevisiae chromosome V
A:Reference number: S53045
A:Accession: S53049
A:Molecule type: DNA
A:Residues: 816-1489 <JAW>
A:Cross-references: EMBL:Z48618; NID:g728690; PID:g728695

R:James, C.M.; Indge, K.J.; Oliver, S.G.
Yeast 11, 1413-1419, 1995
A:Title: DNA sequence analysis of a 35 kb segment from *Saccharomyces cerevisiae* chromosome 11
A:Reference number: S60417; MID:96158061
A:Accession: S60421
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 728-767; 808-862; 865-969; 1001-1021; 1296-1347; 1360-1468 <JAF>
A:Cross-references: EMBL:248618
R:James, C.M.; Indge, K.J.; Oliver, S.G.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64165
A:Accession: S64166
A:Molecule type: DNA
A:Residues: 1-1489 <JAF>
A:Cross-references: EMBL:272672; NID:g1322733; PID:e243594; PID:g1322734; MIPS:YGL150c
A:Experimental source: strain S288C
R:Volckaert, G.; Voet, M.; Verhasselt, P.; Defoor, E.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64153
A:Accession: S64164
A:Molecule type: DNA
A:Residues: 1-137 <VOL>
A:Cross-references: EMBL:272672; MIPS:YGL150c
A:Experimental source: strain S288C
C:Genetics:
A:Map position: 7L

Query Match 96.6%; Score 28; DB 2; Length 1489;
Best Local Similarity 83.3%; Pred. No. 93;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6

DB 1420 YRLLR 1425

RESULT 3
S50445
hypothetical protein YEL014c - yeast (*Saccharomyces cerevisiae*)
C:Species: *Saccharomyces cerevisiae*
C:Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 26-Aug-1999
A:Accession: S50445
R:Dierich, F.S.
submitted to the EMBL Data Library, December 1994
A:Description: *Saccharomyces cerevisiae* chromosome V cosmids 9871, 8199, 9867, 9495 and
A:Reference number: S50428
A:Accession: S50445
A:Molecule type: DNA
A:Residues: 1-101 <DIE>
A:Cross-references: EMBL:U18530; NID:g602367; PID:g602381; MIPS:YEL014c
C:Genetics:
A:Map position: 5L
C:Superfamily: *Saccharomyces* hypothetical protein YEL014c

Query Match 93.1%; Score 27; DB 2; Length 101;
Best Local Similarity 83.3%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6

DB 92 YRLLR 97

RESULT 4
H75054
hypothetical protein PAB2372 - *Pyrococcus abyssi* (strain Orsay)
C:Species: *Pyrococcus abyssi*
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
A:Accession: H75054
R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999
A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome 8
A:Reference number: A75001
A:Accession: H75054
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-157 <KAW>
A:Cross-references: GB:A7248287; GB:AL096836; NID:g5458657; PIDN:CAB50333.1; PID:el51
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB2372

Query Match 93.1%; Score 27; DB 2; Length 157;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6

DB 137 YRLLR 142

RESULT 5
S56263
hypothetical protein YFR008w - yeast (*Saccharomyces cerevisiae*)
C:Species: *Saccharomyces cerevisiae*
C:Date: 02-Sep-1995 #sequence_revision 19-Oct-1995 #text_change 05-Dec-1997
C:Accession: S56263
R:Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sas
submitted to the EMBL Data Library, May 1995
A:Description: Analysis of the nucleotide sequence of chromosome VI from *Saccharomyces*
A:Reference number: S56186
A:Accession: S56263
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-221 <MUR>
A:Cross-references: EMBL:D50617; NID:g836685; PID:d1009888; PID:g836763; MIPS:YFR008w
C:Genetics:
A:Map position: 6R

Query Match 93.1%; Score 27; DB 2; Length 221;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6

DB 211 YRLLR 216

RESULT 6
S57158
hypothetical protein YJR135c - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein J2122
C:Species: *Saccharomyces cerevisiae*
C:Date: 08-Jul-1995 #sequence_revision 08-Sep-1995 #text_change 06-Feb-1998
A:Accession: S57158
R:Rose, M.; Koetter, P.; Entian, K.D.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56848
A:Accession: S57158
A:Molecule type: DNA
A:Residues: 1-239 <ROS>
A:Cross-references: EMBL:249635; NID:g1015871; PID:g1015872; MIPS:YJR135c
C:Genetics:
A:Gene: SGD:MCM22
A:Cross-references: SGD:S0003896; MIPS:YJR135c
A:Map position: 10R

Query Match 93.1%; Score 27; DB 2; Length 239;
Best Local Similarity 83.3%; Pred. No. 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6
||||:|
Db 205 YRLLR 210

RESULT 7
JC4001
macrolide 3-O-acyltransferase (EC 2.3.1.-) - Streptomyces sp.
C:Species: Streptomyces sp.
C:Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 18-Jun-1999
C:Accession: JC4001
R:Arisawa, A.; Tsunekawa, H.; Okamura, K.; Okamoto, R.
Biosci. Biotechnol. Biochem. 59, 582-588, 1995
A:Title: Nucleotide sequence analysis of the carbomycin biosynthetic genes including the
A:Reference number: JC4001; MUID:95290751
A:Accession: JC4001
A:Molecule type: DNA
A:Residues: 1-389 <ARI>
A:Cross-references: DBJ:D30759; NID:q551628; PIDN:BAA06421.1; PID:dl006991; PID:g551631
A:Note: the source was designated as Streptomyces thermotolerans
C:Genetics:
A:Gene: acyA
C:Superfamily: 4''-O-acyltransferase
C:Keywords: acyltransferase

Query Match 93.1%; Score 27; DB 2; Length 389;
Best Local Similarity 83.3%; Pred. No. 44;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6
||||:|
Db 158 YRLLR 163

RESULT 8
D71961
glucose inhibited division protein A - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 26-Aug-1999
C:Accession: D71961
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557
A:Accession: D71961
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-621 <ARN>
A:Cross-references: GB:AE001458; GB:AE001439; NID:g4154713; PIDN:AA005783.1; PID:g415472
A:Experimental source: strain J99
C:Genetics:
A:Gene: gida
A:Superfamily: gida protein

Query Match 93.1%; Score 27; DB 2; Length 621;
Best Local Similarity 83.3%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6
||||:|
Db 432 YRLLR 437

RESULT 9
E64546
glucose inhibited division protein - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 26-Aug-1999

C:Accession: E64546
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKee
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467
A:Accession: E64546
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-621 <TON>
A:Cross-references: GB:AE000541; GB:AE000511; NID:g2313299; PIDN:AA007281.1; PID:g2313299
C:Genetics:
A:Start codon: GTG
C:Superfamily: gida protein

Query Match 93.1%; Score 27; DB 2; Length 621;
Best Local Similarity 83.3%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6
||||:|
Db 432 YRLLR 437

RESULT 10
BWBSGA
glucose-inhibited division protein gida - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 16-Jul-1999
C:Accession: I40440; S66025; G69631; JQ1216; S18076
R:Ogasawara, N.; Yoshikawa, H.
Mol. Microbiol. 6, 629-634, 1992
A:Title: Genes and their organization in the replication origin region of the bacteri
A:Reference number: I40435; MUID:92204018
A:Accession: I40440
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-628 <RES>
A:Cross-references: EMBL:X62539; NID:g40020; PIDN:CAA44404.1; PID:g40026
A:Experimental source: strain CRK2000
R:Ogasawara, N.; Nakai, S.; Yoshikawa, H.
DNA Res. 1, 1-14, 1994
A:Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis ch
A:Reference number: S65967; MUID:96051385
A:Accession: S66025
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-628 <OGN>
A:Cross-references: EMBL:D26185; NID:g467326; PIDN:BAA05231.1; PID:dl005773; PID:g467
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwel, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmer, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati,
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koeter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil
A:Reference number: A69580; MUID:98044033
A:Accession: G69631
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-628 <KUN>
A:Cross-references: GB:Z99124; GB:AL009126; NID:g2636442; PID:el184827; PID:g2636648

A:Experimental source: strain 168

C:Genetics:

A:Gene: gida

C:Superfamily: gida protein

Query Match 93.1%; Score 27; DB 1; Length 628;

Best Local Similarity 83.3%; Pred. No. 71;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6

Db 435 YRLLR 440

RESULT 11

I64078

gida protein - Haemophilus influenzae (strain Rd KW20)

N:Alternate names: glucose-inhibited division protein

C:Species: Haemophilus influenzae

C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 26-Aug-1999

C:Accession: I64078

R:Flerschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.

; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350630

A:Accession: I64078

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-629 <TIGR>

A:Cross-references: GB:U32739; GB:L42023; NID:gl5735559; PIDN:AAC22240.1; PID:gl573570;

C:Genetics:

A:Gene: gida

C:Superfamily: gida protein

Query Match

Best Local Similarity 93.1%; Score 27; DB 2; Length 629;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6

Db 435 YRLLR 440

RESULT 12

F72400

glucose-inhibited division protein A - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 23-Jul-1999

C:Accession: F72400

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Dodson, R.J.; Haft, D.H.; Hickey

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.

C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A:Reference number: A72200; MUID:99287316

A:Accession: F72400

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-629 <ARN>

A:Cross-references: GB:AE001708; GB:AE000512; NID:g4980740; PID:g4980759; TIGR:TM0263

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM0263

C:Superfamily: gida protein

Query Match

Best Local Similarity 93.1%; Score 27; DB 2; Length 629;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6

Db 745 YRLLR 750

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6

Db 442 YRLLR 447

RESULT 13

A48723

replication licensing factor MCM5 - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: cell division control protein CDC46; cell division control protein

C:Species: Schizosaccharomyces pombe

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 13-Aug-1999

C:Accession: A48723

R:Miyake, S.; Okishio, N.; Samejima, I.; Hiraoka, Y.; Toda, T.; Saitoh, I.; Yanagida,

Mol. Biol. Cell 4, 1003-1015, 1993

A:Title: Fission yeast genes nda1(+) and nda4(+), mutations of which lead to S-phase

A:Reference number: A48723; MUID:94129084

A:Accession: A48723

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-720 <MIY>

A:Cross-references: GB:S68467; NID:9545210; PIDN:AAC05058.1; PID:9545211

C:Comment: The complex of six MCM proteins is one of several proteins that must be bo

phosphorylated and dissociate from the chromatin.

C:Genetics:

A:Gene: nda4+

C:Complex: The predominant form is a heterohexamer of MCM2, MCM3, MCM4, MCM5, MCM6, a

C:Function:

A:Description: part of the replication licensing system that permits DNA replication

C:Superfamily: replication licensing factor MCM5; MCM homology

C:Keywords: cell cycle control; DNA replication initiation; nucleus

Query Match

Best Local Similarity 93.1%; Score 27; DB 1; Length 720;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6

Db 675 YRLLR 680

RESULT 14

T10748

mannan endo-1,4-beta-mannosidase (EC 3.2.1.78) - Rhodothermus marinus

N:Alternate names: endo-1,4-beta-mannanase

C:Species: Rhodothermus marinus

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T10748

R:Politz, O.; Krah, M.; Borriess, R.

Submitted to the EMBL Data Library, February 1998

A:Description: Cloning, sequencing and characterization of a beta-mannanase gene of R

A:Reference number: Z17115

A:Accession: T10748

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1021 <POL>

A:Cross-references: EMBL:X90947; NID:el249550; PID:el249551

A:Experimental source: strain ATCC 43812

C:Genetics:

A:Gene: manA

C:Keywords: glycosidase; hydrolase

Query Match

Best Local Similarity 93.1%; Score 27; DB 2; Length 1021;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6

Db 745 YRLLR 750

RESULT 15

S48244
NMD2 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: IFS1 protein; protein YHR077c
C:Species: Saccharomyces cerevisiae
C:Date: 28-Oct-1994 #sequence.revision 10-Feb-1995 #text_change 06-Feb-1998
C:Accession: S48244; S46815; S64648; S64738
R:He, F.; Jacobson, A.
submitted to the EMBL Data Library, September 1994
A:Description: Identification of a novel component of the nonsense-mediated mRNA decay pathway
A:Reference number: S48244
A:Accession: S48244
A:Molecule type: DNA
A:Residues: 1-1089 <HEF>
A:Cross-references: EMBL:U14974; NID:g555938; PID:g555939
R:Favell, T.
submitted to the EMBL Data Library, June 1994
A:Description: The sequence of S. cerevisiae cosmid 9205.
A:Reference number: S46794
A:Accession: S46815
A:Molecule type: DNA
A:Residues: MYQQ',3-1089 <FAV>
A:Cross-references: EMBL:U10556; NID:g500825; PID:g500836; MIPS:YHR077c
R:Lee, S.I.; Umen, J.G.; Varmus, H.E.
Proc. Natl. Acad. Sci. U.S.A. 92, 6587-6591, 1995
A:Title: A genetic screen identifies cellular factors involved in retroviral -1 frameshift
A:Reference number: S64648; MUID:95327692
A:Accession: S64648
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1089 <LEE>
A:Cross-references: EMBL:U28158
R:Varmus, H.E.
submitted to the EMBL Data Library, May 1995
A:Reference number: S64738
A:Accession: S64738
A:Molecule type: DNA
A:Residues: MYQQ',3-1089 <VAR>
A:Cross-references: EMBL:U28158; NID:g967212; PID:g967213
C:Genetics:
A:Gene: SGD:NMD2; IFS1
A:Cross-references: SGD:S0001119; MIPS:YHR077c
A:Map position: 8R
A:Introns: 2/3

Query Match 93.1%; Score 27; DB 2; Length 1089;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6

Db 580 YRLIIR 585

Search completed: February 7, 2000, 11:54:18
Job time: 24328 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 00:59:46 ; Search time 63.71 Seconds
(without alignments)
2.813 Million cell updates/sec

Title: US-08-653-294-6

Perfect score: 29
Sequence: 1 YRLIR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_38.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Query Length | DB ID | Description |
|------------|-------|-------|--------------|--------------|--------------------|
| 1 | 28 | 96.6 | 1489 | 1 YGPO_YEAST | P53115 saccharomyc |
| 2 | 27 | 93.1 | 101 | 1 YEB4_YEAST | P39999 saccharomyc |
| 3 | 27 | 93.1 | 221 | 1 YF88_YEAST | P43592 saccharomyc |
| 4 | 27 | 93.1 | 239 | 1 YJ9D_YEAST | P47167 saccharomyc |
| 5 | 27 | 93.1 | 309 | 1 YF2B_SULSH | P50387 sulfolobus |
| 6 | 27 | 93.1 | 621 | 1 GIDA_HELPU | P56138 helicobacte |
| 7 | 27 | 93.1 | 627 | 1 GIDA_COXBU | P94613 coxiella bu |
| 8 | 27 | 93.1 | 628 | 1 GIDA_BACSU | P25812 bacillus su |
| 9 | 27 | 93.1 | 629 | 1 GIDA_HAEIN | P44763 haemophilus |
| 10 | 27 | 93.1 | 720 | 1 NDA4_SCHPO | P41389 schizosacch |
| 11 | 27 | 93.1 | 1021 | 1 MANA_RHOMR | P49425 rhodothermu |
| 12 | 27 | 93.1 | 1089 | 1 NMD2_YEAST | P38798 saccharomyc |
| 13 | 27 | 93.1 | 1459 | 1 YFIM_CAEEL | Q21874 caenorhabdi |
| 14 | 26 | 89.7 | 293 | 1 RL5_CAEEL | P49405 caenorhabdi |
| 15 | 26 | 89.7 | 299 | 1 RL5_BOMMO | O76190 bombyx mori |
| 16 | 25 | 86.2 | 233 | 1 PRC3_CARAU | O73672 carassius a |
| 17 | 25 | 86.2 | 234 | 1 PRC3_XENLA | P24495 xenopus lae |
| 18 | 25 | 86.2 | 234 | 1 PRC3_DROME | P40301 drosophila |
| 19 | 25 | 86.2 | 259 | 1 LPXA_CHRVI | Q48481 chromatiu |
| 20 | 25 | 86.2 | 271 | 1 RL5_DUNSA | O22608 dunaliella |
| 21 | 25 | 86.2 | 294 | 1 RL5A_SCHPO | P25822 schizosacch |
| 22 | 25 | 86.2 | 294 | 1 RL5B_SCHPO | P47306 schizosacch |
| 23 | 25 | 86.2 | 297 | 1 RL5_HELAN | O65353 helianthus |
| 24 | 25 | 86.2 | 297 | 1 RL5_YEAST | P26321 saccharomyc |
| 25 | 25 | 86.2 | 301 | 1 RL5_NEUCR | O39593 neurospora |
| 26 | 25 | 86.2 | 535 | 1 GIDA_LACLC | O32806 lactococcus |
| 27 | 25 | 86.2 | 629 | 1 GIDA_ECOLI | P17112 escherichia |
| 28 | 25 | 86.2 | 630 | 1 GIDA_PSEPU | P25756 pseudomonas |
| 29 | 25 | 86.2 | 635 | 1 GIDA_SYNY3 | O55694 synechocyst |
| 30 | 25 | 86.2 | 1075 | 1 RPOB_PINTH | P41607 pinus thunb |
| 31 | 25 | 86.2 | 3672 | 1 LML2_CAEEL | Q21313 caenorhabdi |
| 32 | 24 | 82.8 | 81 | 1 YHHP_ECOLI | P37618 escherichia |
| 33 | 24 | 82.8 | 121 | 1 SR14_ARATH | O04421 arabidopsis |
| 34 | 24 | 82.8 | 124 | 1 RL5_PIG | O95276 sus scrofa |

RESULT 1

| ID | YGPO_YEAST | STANDARD; | PRT; 1489 AA. |
|----|---|-----------|------------------|
| AC | P53115; | | |
| DT | 01-OCT-1996 (Rel. 34, Created) | | |
| DT | 01-OCT-1996 (Rel. 34, Last sequence update) | | |
| DT | 01-NOV-1997 (Rel. 35, Last annotation update) | | |
| DE | HYPOTHETICAL 171.5 KD HELICASE IN NUT1-ARO2 INTERGENIC REGION. | | |
| GN | YGL150C OR G1880. | | |
| OS | Saccharomyces cerevisiae (Baker's yeast). | | |
| OC | Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales; | | |
| OC | Saccharomycetaceae; Saccharomycetes. | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RX | MEDLINE; 96158061. | | |
| RA | JAMES C.M., INDGE K.J., OLIVER S.G.; | | |
| RT | "DNA sequence analysis of a 35 kb segment from Saccharomyces cerevisiae chromosome VII reveals 19 open reading frames including RAD54, ACE1/CUP2, PMR1, RCK1, AMS1 and CAL1/CDC43."; | | |
| RT | Yeast 11:1413-1419(1995). | | |
| RL | [2] | | |
| RN | SEQUENCE FROM N.A. | | |
| RC | STRAIN-S288C / FY1769; | | |
| RX | MEDLINE; 97197983. | | |
| RA | VOET M., DEFOOR E., VERHASSELT P., RILES L., ROBBEN J., VOLCKAERT G.; | | |
| RT | "The sequence of a nearly unclonable 22.8 kb segment on the left arm chromosome VII from Saccharomyces cerevisiae reveals ARO2, RPL9A, TIPI, MRF1 genes and six new open reading frames."; | | |
| RT | Yeast 13:177-182(1997) | | |
| RL | [1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL). | | |
| CC | !- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY. | | |
| CC | ----- | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | | |
| CC | ----- | | |
| CC | EMBL; Z48618; CAA88537.1; - | | |
| DR | EMBL; Z72872; CAA96861.1; - | | |
| DR | EMBL; X99960; CAA68224.1; - | | |
| DR | PFAM; PF00176; SNF2_N; 1. | | |
| DR | PFAM; PF00271; helicase.C; 1. | | |
| KW | Hypothetical protein; Nuclear protein; DNA-binding; Helicase; | | |
| KW | ATP-binding. | | |
| FT | DOMAIN 188 193 | | POLY-ALA. |
| FT | DOMAIN 259 268 | | POLY-GLU. |
| FT | DOMAIN 300 306 | | POLY-SER. |
| FT | DOMAIN 568 573 | | POLY-GLU. |
| FT | DOMAIN 675 682 | | POLY-GLU. |
| FT | DOMAIN 731 738 | | ATP (POTENTIAL). |
| FT | NP_BIND 841 844 | | DEAQ BOX. |
| FT | SITE | | |
| FT | SEQUENCE 1489 AA; 171454 MW; 8149887E CRC32; | | |

| | | | | | | |
|----|----|------|-----|---|-------------|--------------------|
| 35 | 24 | 82.8 | 156 | 1 | Y330_RICPR | O92dj8 rickettsia |
| 36 | 24 | 82.8 | 158 | 1 | V56_HPV18 | P06463 human papil |
| 37 | 24 | 82.8 | 158 | 1 | V56_HPV39 | P24835 human papil |
| 38 | 24 | 82.8 | 158 | 1 | V56_HPV45 | P21735 human papil |
| 39 | 24 | 82.8 | 158 | 1 | V56_HPV68 | P54667 human papil |
| 40 | 24 | 82.8 | 181 | 1 | VC16_VACCC | P21100 vaccinia vi |
| 41 | 24 | 82.8 | 183 | 1 | RL18_CICAC | O65729 cicar arlet |
| 42 | 24 | 82.8 | 187 | 1 | RL18_ARATH | P42791 arabidopsis |
| 43 | 24 | 82.8 | 201 | 1 | YEL18_HAEIN | P44189 haemophilus |
| 44 | 24 | 82.8 | 205 | 1 | EPIQ_STAEP | P30198 staphylococ |
| 45 | 24 | 82.8 | 258 | 1 | YDF2_SCHPO | Q10474 schizosacch |

ALIGNMENTS

Query Match 96.6%; Score 28; DB 1; Length 1489;
 Best Local Similarity 83.3%; Pred. No. 49;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YRLLR 6

Db 1420 YRLLR 1425

RESULT 2

YEB4_YEAST STANDARD; PRT; 101 AA.
 AC P39999;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HYPOTHETICAL 12.2 KD PROTEIN IN PMP2-VAC8 INTERGENIC REGION.
 GN YEL014C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
 CC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA DIETRICH F.S., MULLIGAN J.T., HENNESSEY K.M., ALLEN E., ARAUJO R.,
 RA AVILES E., BERNO A., BRENNAN T., CARPENTER J., CHEN E., CHERRY J.M.,
 RA CHUNG E., DUNCAN M., GUZMAN E., HARTZELL G., HUNCKE-SMITH S.,
 RA HYMAN R., KAYSER A., KOMP C., LASHKARI D., LEW H., LIN D.,
 RA MOSEDALE D., NAKAHARA K., NAMATH A., NORGREN R., OEFNER P., OH C.,
 RA PETEL F.X., ROBERTS D., SEHL P., SCHRAMM S., SHOGREN T., SMITH V.,
 RA TAYLOR P., WEI Y., YELTON M., BOYSTEIN D., DAVIS R.W.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U18530; AAB64491.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 101 AA; 12209 MW; BE5A484B CRC32;

Query Match 93.1%; Score 27; DB 1; Length 101;
 Best Local Similarity 83.3%; Pred. No. 5.3;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YRLLR 6

Db 92 YRLLR 97

RESULT 3

YF8_YEAST STANDARD; PRT; 221 AA.
 AC P43592;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE HYPOTHETICAL 25.9 KD PROTEIN IN MPRI-GCN20 INTERGENIC REGION.
 GN YFR008W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
 CC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RX MEDLINE; 95400292.
 RA MURAKAMI Y., NAITOU M., HAGIWARA H., SHIBATA T., OZAWA M.,
 RA SASANUMA S.-I., SASANUMA M., TSUCHIYA Y., SOEDA E., YOKOYAMA K.,

RA YAMAZAKI M., TASHIRO H., EKI T.;
 RT "Analysis of the nucleotide sequence of chromosome VI from
 RT Saccharomyces cerevisiae";
 RL Nat. Genet. 10:261-268(1995).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D50617; BAA09247.1; -;
 KW Hypothetical protein.
 FT DOMAIN 11 22 POLY-GLN.
 SQ SEQUENCE 221 AA; 25915 MW; 3D34764A CRC32;

Query Match 93.1%; Score 27; DB 1; Length 221;
 Best Local Similarity 83.3%; Pred. No. 12;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YRLLR 6

Db 211 YRLLR 216

RESULT 4

YJ9D_YEAST STANDARD; PRT; 239 AA.
 AC P47167;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DE HYPOTHETICAL 27.6 KD PROTEIN IN NMD5-HOM6 INTERGENIC REGION.
 GN YJRL35C OR J2122.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
 CC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA ROSE M., KOETTER P., ENTIAN K.D.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z49635; CAA89666.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 239 AA; 27567 MW; 0BF23C6E CRC32;

Query Match 93.1%; Score 27; DB 1; Length 239;
 Best Local Similarity 83.3%; Pred. No. 13;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YRLLR 6

Db 205 YRLLR 210

RESULT 5

TF2B_SULSH STANDARD; PRT; 309 AA.
 ID TF2B_SULSH
 AC P30387;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

```
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE TRANSCRIPTION INITIATION FACTOR IIB HOMOLOG (TFIIB).
GN TFB
OS Sulfolobus shibatae.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobus.
RN [1]
RP SEQUENCE FROM N.A.
RA QURESHI S.A., KHOO B., BAUMANN P., JACKSON S.P.;
RX MEDLINE; 95320218.
RT "Molecular cloning of the transcription factor TFIIB homolog from
Sulfolobus shibatae."
RL Proc. Natl. Acad. Sci. U.S.A. 92:6077-6081(1995).
CC -!- FUNCTION: STABILIZES TBP BINDING TO AN ARCHAEAL BOX-A PROMOTER.
CC ALSO RESPONSIBLE FOR RECRUITING RNA POLYMERASE II TO THE PRE-
CC INITIATION COMPLEX (DNA-TBP-TFIIB) (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TFIIB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U20899; AAA81380.1; -.
DR HSSP; P29095; IATS.
DR PROSITE; PS00782; TFIIB; 2.
DR PFAM; PF00382; transcript_fac2; 2.
KW Transcription regulation; Repeat.
FT REPEAT 130 206
FT REPEAT 224 300
SQ SEQUENCE 309 AA; 34756 MW; 0921E52C CRC32;
Query Match 93.1%; Score 27; DB 1; Length 309;
Best Local Similarity 83.3%; Pred. No. 17;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 YRLLR 6
Db 201 YRLLR 206
RESULT 6
GIDA_HELPY
ID GIDA_HELPY STANDARD; PRT; 621 AA.
AC P56138; O32632;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE GLUCOSE INHIBITED DIVISION PROTEIN A.
GN GIDA OR HP0213.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE; 97394467.
RA TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,
RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,
RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,
RA LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,
RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,
RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,
RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATTHEY L., WALLIN E.,
RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,
RA VENTER J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
pylori."
RL Nature 388:539-547(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=60190;
RA KARITA M., ETTERBECK M.L., FORSYTH M.H., TUMURU M.K.R., BLASER M.J.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NOT KNOWN.
CC -!- SIMILARITY: BELONGS TO THE GIDA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE00541; AAD07281.1; -.
DR EMBL; AF008565; AAB63296.1; -.
DR TIGR; HP0213; -.
DR PROSITE; PS01280; GIDA_1; 1.
DR PROSITE; PS01281; GIDA_2; 1.
DR PFAM; PF01134; GIDA; 1.
FT VARIANT 10 10 V -> I (IN STRAIN 60190).
FT VARIANT 105 105 L -> F (IN STRAIN 60190).
FT VARIANT 193 193 D -> E (IN STRAIN 60190).
FT VARIANT 220 220 T -> A (IN STRAIN 60190).
FT VARIANT 302 302 S -> N (IN STRAIN 60190).
FT VARIANT 353 353 T -> A (IN STRAIN 60190).
FT VARIANT 379 379 A -> D (IN STRAIN 60190).
FT VARIANT 415 415 I -> V (IN STRAIN 60190).
FT VARIANT 457 457 E -> Q (IN STRAIN 60190).
FT VARIANT 480 481 YI -> CV (IN STRAIN 60190).
FT VARIANT 493 493 EV -> KL (IN STRAIN 60190).
FT VARIANT 505 505 D -> N (IN STRAIN 60190).
FT VARIANT 505 505 D -> N (IN STRAIN 60190).
SQ SEQUENCE 621 AA; 69683 MW; 740EAC4 CRC32;
Query Match 93.1%; Score 27; DB 1; Length 621;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 YRLLR 6
Db 432 YRLLR 437
RESULT 7
GIDA_COXBU
ID GIDA_COXBU STANDARD; PRT; 627 AA.
AC P94613;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE GLUCOSE INHIBITED DIVISION PROTEIN A.
GN GIDA.
OS Coxiella burnetii.
OC Bacteria; Proteobacteria; gamma subdivision; Coxiella.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NINE MILE PHASE I;
RA WILLEMS H., JAEGER C.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NOT KNOWN.
CC -!- SIMILARITY: BELONGS TO THE GIDA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```

DR EMBL: Y10436; CAA71459.1; -
DR PROSITE; PS01280; GIDA_1; 1.
DR PROSITE; PS01281; GIDA_2; 1.
DR PFAM; PF01134; GIDA; 1.
SQ SEQUENCE 627 AA; 69951 MW; B9AF4071 CRC32;

Query Match          93.1%; Score 27; DB 1; Length 627;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6
DB 436 YRLLR 441

RESULT 8
GIDA_BACSU STANDARD; PRT; 628 AA.
AC P25812;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 01-NOV-1995 (Rel. 32, Last annotation update)
DE GLUCOSE INHIBITED DIVISION PROTEIN A.
GN GIDA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / CRK2000;
RX MEDLINE: 92204018.
RA OGASAWARA N., YOSHIKAWA H.;
RT "Genes and their organization in the replication origin region of the
RT bacterial chromosome.";
RL Mol. Microbiol. 6:629-634(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE: 96051395.
RA OGASAWARA N., NAKAI S., YOSHIKAWA H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin.";
RL DNA Res. 1:1-14(1994).
CC -!- FUNCTION: NOT KNOWN.
CC -!- SIMILARITY: BELONGS TO THE GIDA FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X62539; CAA44404.1; -
DR EMBL: D26185; BAA05231.1; -
DR EMBL: Z99124; CAB16138.1; -
DR PIR; JQ1216; BWBSGA.
DR SUBTILIST; BG10059; GIDA.
DR PROSITE; PS01280; GIDA_1; 1.
DR PROSITE; PS01281; GIDA_2; 1.
DR PFAM; PF01134; GIDA; 1.
SQ SEQUENCE 628 AA; 69752 MW; F9B082BB CRC32;

Query Match          93.1%; Score 27; DB 1; Length 628;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6
DB 435 YRLLR 440

RESULT 9
GIDA_HAEIN STANDARD; PRT; 629 AA.
AC P44763;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 01-NOV-1995 (Rel. 32, Last annotation update)
DE GLUCOSE INHIBITED DIVISION PROTEIN A.
GN GIDA OR HI0582.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20;
RX MEDLINE: 95350630.
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,
RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODER A., KELLEY J.M.,
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,
RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
RA FINE L.D., FRITZMAN J.L., FUHRMANN J.L., GEOHAGEN N.S.M.,
RA GNEHM C.L., MCDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA VENTER J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: NOT KNOWN.
CC -!- SIMILARITY: BELONGS TO THE GIDA FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32739; AAC22240.1; -
DR TIGR; HI0582;
DR PROSITE; PS01280; GIDA_1; 1.
DR PROSITE; PS01281; GIDA_2; 1.
DR PFAM; PF01134; GIDA; 1.
SQ SEQUENCE 629 AA; 70103 MW; 6FD83D06 CRC32;

Query Match          93.1%; Score 27; DB 1; Length 629;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6
DB 435 YRLLR 440

RESULT 10
NDA4_SCHPO STANDARD; PRT; 720 AA.
ID NDA4_SCHPO
AC P41389;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE CELL DIVISION CONTROL PROTEIN NDA4.
GN NDA4 OR SPAC3F10.01.
OS Schizosaccharomyces pombe (Fission yeast).
OS Eukaryota; Fungi; Ascomycota; Archiascomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
RN [1]
RP SEQUENCE FROM N.A.

```

RX MEDLINE: 94129084.
RA MIYAKE S., OKISHIO N., SAMEJIMA I., HIRAOKA Y., TODA T., SAITOH I.,
RA YANAGIDA M.;
RT "Fission yeast genes nda1+ and nda4+, mutations of which lead to
RT S-phase block, chromatin alteration and Ca2+ suppression, are members
RT of the CDC46/MCM2 family.";
RL Mol. Biol. Cell 4:1003-1015(1993).
RN [2]
RP SEQUENCE OF 72-720 FROM N.A.
RC STRAIN-972;
RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: ROLE IN DNA REPLICATION AND ESSENTIAL FOR VIABILITY.
CC -|- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -|- SIMILARITY: BELONGS TO THE MCM FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: S68467; AAC60568.1; -;
DR EMBL: Z69369; CA93299.1; -;
DR PIR: A48723; A48723.
DR PROSITE: PS00847; MCM_1; 1.
DR PROSITE: PS00051; MCM_2; 1.
DR PRAM: PF00493; MCM; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW DNA replication; Cell cycle; ATP-binding.
FT DOMAIN 307 312 POLY-GLU.
FT DOMAIN 322 529 MCM.
FT NP_BIND 372 379 ATP (POTENTIAL).
FT CONFLICT 461 461 A -> R (IN REF. 1).
SQ SEQUENCE 720 AA; 80099 MW; 29D61458 CRC32;

Query Match 93.1%; Score 27; DB 1; Length 720;
Best Local Similarity 83.3%; Pred. No. 41;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 YRLIIR 6
DB 675 YRLIIR 680
|||:|

RESULT 11
MANA_RHOMR
ID MANA_RHOMR STANDARD; PRT; 1021 AA.
AC P49425; 1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MANNA ENDO-1,4-BETA-MANNOSIDASE (EC 3.2.1.78).
GN MANA.
OS Rhodothermus marinus.
RC Bacteria; Cytophagales; Rhodothermus group; Rhodothermus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 43812;
RA POLITZ O., KRAH M., BORRISS R.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: RANDOM HYDROLYSIS OF 1,4-BETA-D-
CC LINKAGES IN MANNANS, GALACTOMANNANS, GLUCOMANNANS, AND
CC GALACTOGLUCOMANNANS.
CC -|- SIMILARITY: BELONGS TO FAMILY 26 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X90947; CAA62442.1; -;
DR Hydrolase; Glycosidase.
KW SEQUENCE 1021 AA; 115791 MW; DD4D2403 CRC32;
SQ

Query Match 93.1%; Score 27; DB 1; Length 1021;
Best Local Similarity 83.3%; Pred. No. 59;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 YRLIIR 6
DB 745 YRLIIR 750
|||:|

RESULT 12
NMD2_YEAST
ID NMD2_YEAST STANDARD; PRT; 1089 AA.
AC P38738;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE NONSENSE-MEDIATED MRNA DECAY PROTEIN 2 (UP-FRAMESHIFT SUPPRESSOR 2).
GN NMD2 OR UPE2 OR IFS1 OR SVAL OR YHR077C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95189083.
RA HE F., JACOBSON A.;
RT Identification of a novel component of the nonsense-mediated mRNA
RT decay pathway by use of an interacting protein screen.";
RL Genes Dev. 9:437-454(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-PLY136;
RX MEDLINE: 95189082.
RA CUI Y., HAGAN K.W., ZHANG S., PELTZ S.W.;
RT Identification and characterization of genes that are required for
RT the accelerated degradation of mRNAs containing a premature
RT translational termination codon.";
RL Genes Dev. 9:423-436(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95327692.
RA LEE S.I., UMEN J.G., VARMUS H.E.;
RT "A genetic screen identifies cellular factors involved in retroviral
RT -1 frameshifting.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:6587-6591(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RX MEDLINE: 94378003.
RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DOVER J.,
RA DU Z., FAVELLO A., FULTON L., GATTUNG S., GEISEL C., KIRSTEN J.,
RA KUCABA T., HILLIER L., JIER M., JOHNSTON L., LANGSTON Y.,
RA LATREILLE P., LOUIS E.J., MACRI C., MARDIS E., MENEZES S., MOUSER L.,
RA NHAN M., RIFKIN L., RILES L., ST PETER H., TREVASKIS E., VAUGHAN K.,
RA VIGNATI D., WILCOX L., WOHLDMAN P., WATERSTON R., WILSON R.,
RA VAUDIN M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VII.";
RL Science 265:2077-2082(1994).
CC -|- FUNCTION: INVOLVED IN NONSENSE-MEDIATED DECAY OF MRNAS CONTAINING
CC PREMATURE STOP CODONS. IT INTERACTS VIA ITS C-TERMINUS, WITH
CC NAM7/UPF1. COULD BE INVOLVED IN DETERMINING THE EFFICIENCY OF
CC TRANSLATIONAL TERMINATION OR REINITIATION OR FACTORS INVOLVED IN
CC THE INITIAL ASSEMBLY OF AN INITIATION- AND TERMINATION-COMPETENT
CC MRNP.

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U14974; AAA67724.1; -
 CC EMBL; U12137; AAA66521.1; -
 CC EMBL; U28158; AAA74948.1; -
 CC EMBL; U10556; AAB68893.1; -
 CC PIR; S46815; S46815.
 CC SGD; L0001257; NMD2.
 CC FT DOMAIN 843 975 ASP/GLU-RICH (HIGHLY ACIDIC).
 CC FT CONFLICT 2 2 D -> YQQ (IN REF. 3 AND 4).
 CC SQ SEQUENCE 1089 AA: 126746 MW: 486488F CRC32;

Query Match 93.1%; Score 27; DB 1; Length 1089;
 Best Local Similarity 83.3%; Pred. No. 63;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLILR 6

Db 580 YRLILR 585

RESULT 13

YEIM_CAEEL STANDARD; PRT; 1459 AA.
 ID YFIM_CAEEL
 AC Q21874;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HYPOTHETICAL 166.0 KD PROTEIN R09E10.5 IN CHROMOSOME IV.
 GN R09E10.5
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
 CC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
 CC [1]
 CC STRAIN-BRISTOL N2;
 CC RA MATTHEWS L.;
 CC RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: STRONG, TO C.ELEGANS F54D1.6.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; Z70287; CAA94300.1; -
 CC DR WORMPEP; R09E10.5; CE06287.
 CC KW Hypothetical protein; Transmembrane.
 CC SQ SEQUENCE 1459 AA: 165994 MW: 46E17445 CRC32;

Query Match 93.1%; Score 27; DB 1; Length 1459;
 Best Local Similarity 83.3%; Pred. No. 85;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLILR 6

Db 188 YRLILR 193

RESULT 14

RL5_CAEEL STANDARD; PRT; 293 AA.
 ID P49105;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE PROBABLE 60S RIBOSOMAL PROTEIN L5.
 GN F54C9.5
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
 CC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
 CC [1]
 CC STRAIN-BRISTOL N2;
 CC RC STRAIN-BRISTOL N2;
 CC RA SIMS M.;
 CC RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: THIS PROTEIN BINDS 5S RNA (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z49967; CAA90251.1; -
 CC DR WORMPEP; F54C9.5; CE02255.
 CC DR PFAM; PF00861; Ribosomal_L18p; 1.
 CC KW Ribosomal protein; rRNA-binding.
 CC SQ SEQUENCE 293 AA: 33386 MW: 8FPEC7EE CRC32;

Query Match 89.7%; Score 26; DB 1; Length 293;
 Best Local Similarity 66.7%; Pred. No. 28;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLILR 6

Db 49 YRLILR 54

RESULT 15

RL5_BOMMO STANDARD; PRT; 299 AA.
 ID RL5_BOMMO
 AC O76190;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE 60S RIBOSOMAL PROTEIN L5.
 GN RPL5.
 OS Bombyx mori (Silk moth).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Eukaryota; Metazoa; Arthropoda; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 CC OC Pterygota; Neoptera; Bombycoidea; Bombycidae; Bombyx.
 CC [1]
 CC SEQUENCE FROM N.A.
 CC RC TISSUE-SILK GLAND;
 CC RA YANG C.S., SEHNAL F.;
 CC RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: THIS PROTEIN BINDS 5S RNA (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; AF008229; AAC24960.1; -

DR PFAM; PF00861; Ribosomal_L18p; 1.
KW Ribosomal protein; rRNA-binding.
SQ SEQUENCE 299 AA; 34378 MW; 7262D2FC CRC32;

Query Match 89.7%; Score 26; DB 1; Length 299;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 YRLIIR 6
| | | : |
Db 49 YRLIVR 54

Search completed: February 8, 2000, 00:59:48
Job time: 377 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 13:17:31 ; Search time 209.03 Seconds
(without alignments)
1.990 Million cell updates/sec

Title: US-08-653-294-6

Perfect score: 29

Sequence: 1 YRLIIR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL_12:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 29 | 100.0 | 2692 | 5 Q21547 | Q21547 caenorhabdi |
| 2 | 28 | 96.6 | 325 | 1 Q9Y942 | Q9Y942 aeropyrum p |
| 3 | 27 | 93.1 | 389 | 2 Q56074 | Q56074 streptomyce |
| 4 | 27 | 93.1 | 418 | 2 Q9X8L1 | Q9X8L1 streptomyce |
| 5 | 27 | 93.1 | 621 | 2 Q9ZML9 | Q9ZML9 helicobacte |
| 6 | 27 | 93.1 | 629 | 2 Q9WYAL | Q9WYAL thermotoga |
| 7 | 27 | 93.1 | 724 | 2 Q45730 | Q45730 bacillus th |
| 8 | 27 | 93.1 | 1054 | 4 Q94887 | Q94887 homo sapien |
| 9 | 26 | 89.7 | 116 | 1 Q9YF18 | Q9YF18 aeropyrum p |
| 10 | 26 | 89.7 | 173 | 2 Q51316 | Q51316 nostoc sp. |
| 11 | 26 | 89.7 | 233 | 1 Q26858 | Q26858 methanobact |
| 12 | 26 | 89.7 | 296 | 5 Q9Y0H6 | Q9Y0H6 myxine glut |
| 13 | 26 | 89.7 | 357 | 2 Q9ZK14 | Q9ZK14 helicobacte |
| 14 | 26 | 89.7 | 374 | 2 Q83496 | Q83496 treponema p |
| 15 | 26 | 89.7 | 447 | 5 Q21147 | Q21147 caenorhabdi |
| 16 | 26 | 89.7 | 469 | 2 Q83593 | Q83593 treponema p |
| 17 | 26 | 89.7 | 750 | 2 Q9Z1U5 | Q9Z1U5 bacillus th |
| 18 | 26 | 89.7 | 770 | 11 Q61139 | Q61139 mus musculu |
| 19 | 26 | 89.7 | 783 | 11 Q62849 | Q62849 rattus norv |
| 20 | 26 | 89.7 | 785 | 4 Q16549 | Q16549 homo sapien |

| | | | | | |
|----|----|------|------|-----------|--------------------|
| 21 | 26 | 89.7 | 2263 | 12 Q82998 | Q82998 bunyavirus |
| 22 | 25 | 86.2 | 145 | 2 Q34084 | Q34084 streptococc |
| 23 | 25 | 86.2 | 191 | 12 Q98543 | Q98543 paramecium |
| 24 | 25 | 86.2 | 380 | 1 Q9YDM1 | Q9YDM1 aeropyrum p |
| 25 | 25 | 86.2 | 415 | 4 Q95906 | Q95906 homo sapien |
| 26 | 25 | 86.2 | 416 | 11 Q62083 | Q62083 mus musculu |
| 27 | 25 | 86.2 | 445 | 5 Q22419 | Q22419 caenorhabdi |
| 28 | 25 | 86.2 | 481 | 2 Q9XB67 | Q9XB67 bacteroides |
| 29 | 25 | 86.2 | 558 | 5 Q93329 | Q93329 caenorhabdi |
| 30 | 25 | 86.2 | 605 | 5 Q93607 | Q93607 treponema p |
| 31 | 25 | 86.2 | 622 | 5 P90866 | P90866 caenorhabdi |
| 32 | 25 | 86.2 | 1264 | 5 P91767 | P91767 manduca sex |
| 33 | 25 | 86.2 | 1846 | 5 Q61776 | Q61776 caenorhabdi |
| 34 | 25 | 86.2 | 2228 | 10 Q48579 | Q48579 arabidopsis |
| 35 | 25 | 86.2 | 3704 | 5 P91904 | P91904 caenorhabdi |
| 36 | 24 | 82.8 | 63 | 12 Q98229 | Q98229 molluscum c |
| 37 | 24 | 82.8 | 63 | 12 Q89560 | Q89560 vaccinia vi |
| 38 | 24 | 82.8 | 63 | 12 Q07047 | Q07047 variola vir |
| 39 | 24 | 82.8 | 87 | 2 Q33348 | Q33348 mycobacteri |
| 40 | 24 | 82.8 | 96 | 2 Q47052 | Q47052 escherichia |
| 41 | 24 | 82.8 | 99 | 2 Q9X2S9 | Q9X2S9 escherichia |
| 42 | 24 | 82.8 | 104 | 12 Q56838 | Q56838 vibrio chol |
| 43 | 24 | 82.8 | 109 | 12 Q11376 | Q11376 molluscum c |
| 44 | 24 | 82.8 | 110 | 12 Q98316 | Q98316 molluscum c |
| 45 | 24 | 82.8 | 141 | 5 Q00825 | Q00825 trypanosoma |

ALIGNMENTS

RESULT 1

Q21547 PRELIMINARY; PRT: 2692 AA.
AC Q21547; Q22086;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE T01H10.8 PROTEIN.
GN T01H10.8
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA MATTHEWS P.
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z67995; CAA91946.1; -
DR EMBL; Z67737; CAA91946.1; JOINED.
DR EMBL; Z67737; CAA91543.1; -
DR EMBL; Z67995; CAA91543.1; JOINED.
SQ SEQUENCE 2692 AA; 312318 MW; 9F782AC7 CRC32;

Query Match 100.0%; Score 29; DB 5; Length 2692;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YRLIIR 6
|||||

Db 1079 YRLIIR 1084

RESULT 2

Q9Y942 PRELIMINARY; PRT: 325 AA.
ID Q9Y942
AC Q9Y942;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE 325AA LONG HYPOTHETICAL TRANSCRIPTION INITIATION FACTOR IIB.
GN APE2443.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Aeropyrum.

```

RN RP SEQUENCE FROM N.A.
RC STRAIN-K1;
RX MEDLINE; 99310339.
RA KAWARABAYASI Y., HINO Y., HORIKAWA H., YAMAZAKI S., HAIKAWA Y.,
RA JIN-NO K., TAKAHASHI M., SEKINE M., BABA S., ANKAI A., KOSUGI H.,
RA HOSOTAWA A., FUKUI S., NAGAI Y., NISHIJIMA K., NAKAZAWA H.,
RA TAKAMIYA M., MASUDA S., FUNAHASHI T., TANAKA T., KUDOH Y.,
RA YAMAZAKI J., KUSHIDA N., OGUCHI A., AOKI K., KUBOTA K., NAKAMURA Y.,
RA NOMURA N., SAKO Y., KIKUCHI H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL: APO00064; BAA81458.1; -.
KW Initiation factor.
SQ SEQUENCE 325 AA; 35735 MW; 4A3FC0C7 CRC32;

Query Match 96.6%; Score 28; DB 1; Length 325;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6
DB 209 YRLDVR 214

RESULT 3
Q56074 PRELIMINARY; PRT; 389 AA.
AC Q56074;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE MACROLIDE ANTIBIOTICS 3-O-ACYLTRANSFERASE.
GN ACYA.
OS Streptomyces thermotolerans.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN RP SEQUENCE FROM N.A.
RC STRAIN-TH475;
RX MEDLINE; 95290751.
RA ARISAWA A., TSUNEKAWA H., OKAMURA K., OKAMOTO R.;
RT "Nucleotide sequence analysis of the carbomycin biosynthetic genes
RT including the 3-O-acyltransferase gene from Streptomyces
RT thermotolerans.";
RL Biosci. Biotechnol. Biochem. 59:582-588(1995).
RN RP SEQUENCE FROM N.A.
RC STRAIN-TH475;
RX MEDLINE; 94354673.
RA ARISAWA A., KAWAMURA N., TAKEDA K., TSUNEKAWA H., OKAMURA K.,
RA OKAMOTO R.;
RT "Cloning of the macrolide antibiotic biosynthesis gene acyA, which
RT encodes 3-O-acyltransferase, from Streptomyces thermotolerans and its
RT use for direct fermentative production of a hybrid macrolide
RT antibiotic.";
RL Appl. Environ. Microbiol. 60:2657-2660(1994).
DR EMBL: D30759; BAA06421.1; -.
KW Transferase; Acyltransferase.
SQ SEQUENCE 389 AA; 42879 MW; 2282D0F7 CRC32;

Query Match 93.1%; Score 27; DB 2; Length 389;
Best Local Similarity 83.3%; Pred. No. 96;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6
DB 158 YRLDVR 163

```

```

RESULT 4
Q9X8L1 PRELIMINARY; PRT; 418 AA.
ID Q9X8L1;
AC Q9X8L1;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
DE HYPOTHETICAL 44.6 KD PROTEIN.
GN SCE9.39C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX OLIVER K., HARRIS D.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX JAMES K.D., PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE; 97000351.
RA REDENBACH M., KIESER H.M., DENAPAITE D., EICHNER A., CULLUM J.,
RA KINASHI H., HOPWOOD D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL049841; CAB42785.1; -.
KW Hypothetical protein.
SQ SEQUENCE 418 AA; 44560 MW; 9E269715 CRC32;

Query Match 93.1%; Score 27; DB 2; Length 418;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLLR 6
DB 126 YRLDVR 131

RESULT 5
Q92ML9 PRELIMINARY; PRT; 621 AA.
ID Q92ML9;
AC Q92ML9;
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
DE GLUCOSE INHIBITED DIVISION PROTEIN A.
GN GIDA.
OS Helicobacter pylori J99.
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
RN RP SEQUENCE FROM N.A.
RC STRAIN-J99;
RX MEDLINE; 99120557.
RA ALM R.A., LING L.-S.L., MOIR D.T., KING B.L., BROWN E.D., DOIG P.C.,
RA SMITH D.R., NOONAN B., GUILD B.C., DEJONGE B.L., CARMEL G.,
RA TUMMINO P.J., CARUSO A., URTA-NICKELSEN M., MILLS D.M., IYES C.,
RA GIBSON R., MERBERG D., MILLS S.D., JIANG Q., TAYLOR D.E., VOVIS G.F.,
RA TRUST T.J.;
RT "Genomic-sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
DR EMBL: AE001458; RAD05783.1; -.
DR PROSITE; PS01280; GIDA_1; 1.

```

DR PROSITE; PS01281; GIDA_2; 1.
SQ SEQUENCE 621 AA; 69691 MW; 928413A8 CRC32;

Query Match 93.1%; Score 27; DB 2; Length 621;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6
|||:|
DB 432 YRLLR 437

RESULT 6
Q9WYAL PRELIMINARY; PRT; 629 AA.
AC Q9WYAL;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE GLUCOSE-INHIBITED DIVISION PROTEIN A.
GN TM0263.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 9287316.
RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,
RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,
RA MCDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,
RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,
RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,
RA SMITH H.O., VENTER J.C., FRASER C.M.;
RL "Evidence for lateral gene transfer between Archaea and bacteria from
genome sequence of Thermotoga maritima";
RL Nature 399:323-329(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,
RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,
RA MCDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,
RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,
RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,
RA SMITH H.O., VENTER J.C., FRASER C.M.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE01708; AAD35351.1;
DR PROSITE; PS01280; GIDA_1; 1.
DR PROSITE; PS01281; GIDA_2; 1.
SQ SEQUENCE 629 AA; 71039 MW; DA12D9CD CRC32;

Query Match 93.1%; Score 27; DB 2; Length 629;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6
|||:|
DB 442 YRLLR 447

RESULT 7
Q45730 PRELIMINARY; PRT; 724 AA.
AC Q45730;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE MOSQUITOCIDAL TOXIN.
GN CRT11B.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-JEGATHESAN;
RX MEDLINE: 96086008.
RA DELECLUSE A., ROSSO M.L., RAGNI A.;
RT "Cloning and expression of a novel toxin gene from Bacillus
thuringiensis subsp. jegathesan encoding a highly mosquitocidal
RT protein";
RL Appl. Environ. Microbiol. 61:4230-4250(1995).
DR EMBL; X86902; CAA60504.1; -;
SQ SEQUENCE 724 AA; 81343 MW; 8BB449D1 CRC32;

Query Match 93.1%; Score 27; DB 2; Length 724;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6
|||:|
DB 548 YRLIR 553

RESULT 8
O94887 PRELIMINARY; PRT; 1054 AA.
AC O94887;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE KIAA0793 PROTEIN.
GN KIAA0793.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE: 99087487.
RA NAGASE T., ISHIKAWA K., SUYAMA M., KIKUNO R., MIYAJIMA N., TANAKA A.,
RA KOTANI H., NOMURA N., OHARA O.;
RT "Prediction of the coding sequences of unidentified human genes. XI.
The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro";
RL DNA Res. 5:277-286(1998).
RN [2]
RP SEQUENCE OF 1-754 FROM N.A.
RX MEDLINE: 99063792.
RA SULSTON J.E., WATERSTON R.;
RT "Toward a complete human genome sequence.";
RL Genome Res. 8:1097-1108(1998).
RN [3]
RP SEQUENCE OF 1-754 FROM N.A.
RA SUN H., STONEKING T., LANGSTON Y., LAPLANT Y.;
RT "The sequence of Homo sapiens BAC clone RG442F18.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-754 FROM N.A.
RA WATERSTON R.H.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-754 FROM N.A.
RA WATERSTON R.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE OF 1-754 FROM N.A.
RA WATERSTON R.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB018336; BAA34513.1; -;
DR EMBL; AC005104; AAD12224.1; -;
DR HSP; P08567; 1PLS.
SQ SEQUENCE 1054 AA; 119888 MW; 3BA89171 CRC32;

Query Match 93.1%; Score 27; DB 4; Length 1054;

Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6
|:|:|:|
DB 688 YRLIIR 693

RESULT 9
Q9YF18 PRELIMINARY; PRT; 116 AA.
AC Q9YF18;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE 116AA LONG HYPOTHETICAL PROTEIN.
GN APE0420.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Aeropyrum.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE; 99310339.
RA KAWARABAYASI Y., HINO Y., HORIKAWA H., YAMAZAKI S., HAIKAWA Y.,
RA JIN-NO K., TAKAHASHI M., SEKINE M., BABA S., ANKAI A., KOSUGI H.,
RA HOSOVAMA A., FUKUI S., NAGAI Y., NISHIJIMA K., NAKAZAWA H.,
RA TAKAMIYA M., MASUDA S., FUNAHASHI T., TANAKA T., KUDOH Y.,
RA YAMAZAKI J., KUSHIDA N., OGUCHI A., AOKI K., KUBOTA K., NAKAMURA Y.,
RA NOMURA N., SAKO Y., KIKUCHI H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000059; BAA79378.1; -
SQ SEQUENCE 116 AA; 12928 MW; 9D725285 CRC32;

Query Match 89.7%; Score 26; DB 1; Length 116;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6
|:|:|:|
DB 3 YRLIIR 8

RESULT 10
Q51316 PRELIMINARY; PRT; 173 AA.
AC Q51316;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE SS DNA REPLICATING PLASMID ENCODING A REPLICATION-ASSOCIATED PROTEIN
DE (REPA) AND THREE ORFS, COMPLETE CDS (REPA).
OS Nostoc sp.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
RN [1]
RP SEQUENCE FROM N.A.
RA WALTON D.K., GENDEL S.M., ATHERLY A.G.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M81381; AAA25516.1; -
SQ SEQUENCE 173 AA; 19346 MW; B40B691A CRC32;

Query Match 89.7%; Score 26; DB 2; Length 173;
Best Local Similarity 83.3%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6
|:|:|:|
DB 164 YRLIIR 169

RESULT 11
O26858 PRELIMINARY; PRT; 233 AA.
AC O26858;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE ENDONUCLEASE III.
GN MTH764.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE; 98037514.
RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,
RA ALDREDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,
RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTHIER B., QIU D.,
RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,
RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,
RA MCDUGALL S., SHIMER G., GOYAL A., PIETROVSKI S., CHURCH G.M.,
RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT J. Bacteriol. 179:7135-7155(1997).
RL EMBL; AE000855; AAB83267.1; -
DR HSSP; P20625; ZABK.
DR PFAM; PF00730; Endonuclease_3; 1.
KW Endonuclease.
SQ SEQUENCE 233 AA; 27012 MW; 6B518190 CRC32;

Query Match 89.7%; Score 26; DB 1; Length 233;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6
|:|:|:|
DB 46 YRLIIR 51

RESULT 12
Q9YOH6 PRELIMINARY; PRT; 296 AA.
AC Q9YOH6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE RIBOSOMAL PROTEIN L5.
OS Myxine glutinosa (Atlantic hagfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Myxini; Myxiniiformes;
OC Myxiniidae; Myxine.
RN [1]
RP SEQUENCE FROM N.A.
RA WHITE G.P., CUNNINGHAM C.;
RT "The 60S ribosomal protein L5 of Atlantic hagfish (Myxine
RT glutinosa).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152472; AAD37804.1; -
KW Ribosomal protein.
SQ SEQUENCE 296 AA; 34296 MW; C3AB83EC CRC32;

Query Match 89.7%; Score 26; DB 5; Length 296;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLIIR 6
|:|:|:|
DB 49 YRLIIR 54

```

RESULT 13
ID Q92KI4 PRELIMINARY; PRT: 357 AA.
AC Q92KI4;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE INTEGRASE/RECOMBINASE (XERCID FAMILY).
GN JHP0951.
OS Helicobacter pylori J99.
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-J99;
RX MEDLINE; 99120557.
RA ALM R.A., LING L.-S.L., MOIR D.T., KING B.L., BROWN E.D., DOIG P.C.,
RA SMITH D.R., NOONAN B., GUILD B.C., DEJONGE B.L., CARMEL G.,
RA TUMMINO P.J., CARUSO A., URIA-NICKELSEN M., MILLS D.M., IVES C.,
RA GIBSON R., MERBERG D., MILLS S.D., JIANG Q., TAYLOR D.E., VOVIS G.F.,
RA TRUST T.J.;
RT "Genomic-sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
DR EMBL: AE001524; AAD06529.1; -.
DR HSP: P21891; IAP.
SQ SEQUENCE 357 AA; 42379 MW; 931DB706 CRC32;

Query Match 89.7%; Score 26; DB 2; Length 357;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YRLIIR 6
Db 216 YRLLIK 221

RESULT 14
ID O83496 PRELIMINARY; PRT: 374 AA.
AC O83496;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE HYPOTHETICAL 42.1 KD PROTEIN.
GN TP0483.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98332770.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.,
RA DODSON R., GWINN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.,
RA SODERREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.,
RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTTERBACK T.,
RA McDONALD L., ARTIACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.,
RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J., SMITH H.O.,
RA VENTER J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.,
RA DODSON R., GWINN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.,
RA SODERREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.,
RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTTERBACK T.,
RA McDONALD L., ARTIACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.,
RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J., SMITH H.O.,
RA VENTER J.C.;
RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE001224; AAC65472.1; -.

```

```

DR TIGR; TP0483; -.
DR PFAM: PF00041; fn3; 1.
KW Hypothetical protein.
SQ SEQUENCE 374 AA; 42126 MW; D696A45B CRC32;

Query Match 89.7%; Score 26; DB 2; Length 374;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YRLIIR 6
Db 48 YRLVLR 53

RESULT 15
ID Q21147 PRELIMINARY; PRT: 447 AA.
AC Q21147;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE SIMILAR TO EPOXIDE HYDROLASE.
GN K02F3.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX BENTLEY D.;
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: U00052; AAA50712.1; -.
DR PFAM: PF00561; abhydrolase; 1.
KW Hydrolase.
SQ SEQUENCE 447 AA; 52991 MW; 399D289B CRC32;

Query Match 89.7%; Score 26; DB 5; Length 447;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YRLIIR 6
Db 72 YRLVLR 77

Search completed: February 8, 2000, 13:17:33
Job time: 32482 sec

```

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-6 to: GenEmbl.* out_format : pfs
Date: Feb 8, 2000 4:37 PM
About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-O=Cg1_1/USPTO_spool/US08653294/runat_04022000_160701_15779/app_query.fasta.1
-DB=GenEmbl -QPMI-fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -XGAPOP=6.000
-GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELEX=7.000 -START=1 -MATRIX=blosom62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR.SCORE=pct -ALIGN=15 -MODE=LOCAL
-OUTFWT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294
-NCPU=6 -ICPU=3 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-6

Query length: 6

Database: GenEmbl.*

Database sequences: 821193

Database length: 1518192014

Search time (sec): 11370.480000

score_list:

| Sequence | Strd | Orig | ZScore | Escore | Len | Documentation |
|------------------|------|-------|--------|---------|-------|--|
| gb_ov:AF012770 | + | 29.00 | 122.09 | 88.79 | 306 | AF012770 Channa gachua GAC4 cytochrome b (cytb) gene, mitochondrial |
| gb_ov:AF012771 | + | 29.00 | 122.09 | 88.79 | 306 | AF012771 Channa gachua GAC6 cytochrome b (cytb) gene, mitochondrial |
| gb_ov:AF012784 | + | 29.00 | 122.09 | 88.79 | 306 | AF012784 Channa micropeltes MIC1 cytochrome b (cytb) gene, mitochondrial |
| gb_ov:AF012785 | + | 29.00 | 122.09 | 88.79 | 306 | AF012785 Channa micropeltes MIC2 cytochrome b (cytb) gene, mitochondrial |
| gb_ov:AF012786 | + | 29.00 | 122.09 | 88.79 | 306 | AF012786 Channa orientalis ORI1 cytochrome b (cytb) gene, mitochondrial |
| gb_ov:AF012787 | + | 29.00 | 122.09 | 88.79 | 306 | AF012787 Channa orientalis ORI2 cytochrome b (cytb) gene, mitochondrial |
| gb_ov:AF012788 | + | 29.00 | 122.09 | 88.79 | 306 | AF012788 Channa striata STR1 cytochrome b (cytb) gene, mitochondrial |
| gb_ov:CGMTCYB1 | + | 29.00 | 122.09 | 88.79 | 306 | AF012789 Channa striata STR1 cytochrome b (cytb) gene, mitochondrial |
| gb_ov:CGMTCYB2 | + | 29.00 | 122.09 | 88.79 | 306 | AF012790 Channa striata STR2 cytochrome b (cytb) gene, mitochondrial |
| gb_ov:CGMTCYB3 | + | 29.00 | 122.09 | 88.79 | 306 | AF012791 Channa striata STR3 cytochrome b (cytb) gene, mitochondrial |
| gb_ov:CGMTCYB4 | + | 29.00 | 122.09 | 88.79 | 306 | AF012792 Channa striata STR4 cytochrome b (cytb) gene, mitochondrial |
| gb_ov:CGMTCYB5 | + | 29.00 | 122.09 | 88.79 | 306 | AF012793 Channa striata STR5 cytochrome b (cytb) gene, mitochondrial |
| gb_ov:PAU26956 | + | 29.00 | 121.26 | 98.78 | 340 | U26956 Pristigaster aequiloides cytochrome b (cytb) gene, mitochondrial |
| gb_ov:FDU77119 | + | 29.00 | 120.88 | 103.78 | 357 | U77119 Fundulus dispar cytochrome b (cytb) gene, mitochondrial |
| gb_ov:FDU77120 | + | 29.00 | 120.88 | 103.78 | 357 | U77120 Fundulus dispar cytochrome b (cytb) gene, mitochondrial |
| gb_p12:GNS019XS | + | 29.00 | 116.03 | 193.32 | 660 | AL112232 Botrytis cinerea strain 112232 cytochrome b (cytb) gene, mitochondrial |
| gb_srs:G59383 | - | 29.00 | 115.94 | 195.40 | 667 | G59383 SHGC-110718 Human Homo sapiens cytochrome b (cytb) gene, mitochondrial |
| gb_p11:SKU26206 | - | 29.00 | 112.41 | 307.23 | 1043 | U26206 Symbiotaphrina kochii, cytochrome b (cytb) gene, mitochondrial |
| gb_p12:HSNCFI13 | + | 29.00 | 111.68 | 337.66 | 1145 | U00788 Human neutrophil oxidase cytochrome b (cytb) gene, mitochondrial |
| gb_v1:HOBHEGA | + | 29.00 | 110.65 | 385.16 | 1304 | M76373 Human coronavirus hemagglutinin cytochrome b (cytb) gene, mitochondrial |
| gb_ba1:ECU20767 | + | 29.00 | 109.77 | 431.54 | 1459 | U20767 Escherichia coli beta-galactosidase cytochrome b (cytb) gene, mitochondrial |
| gb_ba1:ECOPABC | - | 29.00 | 108.71 | 493.86 | 1867 | M93135 Escherichia coli aminodipropionyl transferase cytochrome b (cytb) gene, mitochondrial |
| gb_p11:HUMNOXF | + | 29.00 | 106.50 | 655.80 | 2206 | M32011 Human neutrophil oxidase cytochrome b (cytb) gene, mitochondrial |
| gb_ba2:AF076153 | + | 29.00 | 104.72 | 824.56 | 2766 | AF076153 Salmonella typhimurium cytochrome b (cytb) gene, mitochondrial |
| gb_ov:AB009568 | - | 29.00 | 104.20 | 881.31 | 2954 | AB009568 Oryzias latipes mtDNA cytochrome b (cytb) gene, mitochondrial |
| gb_ba1:SEN237788 | + | 29.00 | 103.83 | 923.91 | 3095 | AJ237788 Salmonella enterica serovar Senftenberg cytochrome b (cytb) gene, mitochondrial |
| gb_in1:DROMYHB | + | 29.00 | 100.57 | 1448.03 | 4876 | J02788 D.melanogaster myosin heavy chain cytochrome b (cytb) gene, mitochondrial |
| gb_p11:PETNIAA | - | 29.00 | 98.12 | 1954.03 | 6380 | L11563 Petunia hybrida nitrate reductase cytochrome b (cytb) gene, mitochondrial |
| gb_p11:PETNITRED | - | 29.00 | 98.12 | 1954.03 | 6380 | L13691 Petunia hybrida nitrate reductase cytochrome b (cytb) gene, mitochondrial |
| gb_p12:AF071550 | + | 29.00 | 97.23 | 2428.03 | 7145 | AF071550 Vigna radiata (orif140) cytochrome b (cytb) gene, mitochondrial |
| gb_ba1:ECRHSBG | + | 29.00 | 96.61 | 2438.03 | 7721 | X60999 Escherichia coli rhdD gene, mitochondrial |
| gb_ba1:ECORHSDX | + | 29.00 | 96.61 | 2438.03 | 7723 | L19084 Escherichia coli rhdD gene, mitochondrial |
| gb_p11:HSALURPYS | + | 29.00 | 96.14 | 2458.03 | 8200 | U24594 H.sapiens DNA for alu element cytochrome b (cytb) gene, mitochondrial |
| gb_p1:AF066865 | - | 29.00 | 96.12 | 2458.03 | 8219 | AF066865 Bacteriophage TPW22 cytochrome b (cytb) gene, mitochondrial |
| gb_ba2:AE000156 | + | 29.00 | 93.90 | 3438.03 | 11589 | AE000156 Deinococcus radiodurans cytochrome b (cytb) gene, mitochondrial |
| gb_ba2:AE0002079 | + | 29.00 | 93.41 | 3438.03 | 11589 | AE0002079 Deinococcus radiodurans cytochrome b (cytb) gene, mitochondrial |
| gb_ba2:AE000210 | + | 29.00 | 92.47 | 4088.03 | 13051 | AE000210 Escherichia coli K-12 cytochrome b (cytb) gene, mitochondrial |
| gb_ba1:BBGRSB | + | 29.00 | 92.02 | 4288.03 | 13828 | X61658 B.brevis grsB gene for cytochrome b (cytb) gene, mitochondrial |
| gb_ba2:AE001265 | + | 29.00 | 91.99 | 4288.03 | 13865 | AE001265 Treponema pallidum ssp. pallidum cytochrome b (cytb) gene, mitochondrial |
| gb_ba1:BACGS2 | + | 29.00 | 91.89 | 4388.03 | 14048 | D29676 B. brevis grs2 gene encoding cytochrome b (cytb) gene, mitochondrial |
| gb_ba2:AE000704 | + | 29.00 | 91.72 | 4488.03 | 14351 | AE000704 Aquifex aeolicus secY gene, mitochondrial |
| gb_ba1:D90745 | - | 29.00 | 90.68 | 5088.03 | 16379 | D90745 Escherichia coli genome cytochrome b (cytb) gene, mitochondrial |

gb_htg3:AC010371 + 29.00 90.56 5.1e+03 16637 ! AC010371 Homo sapiens chro
gb_in1:CEU26736 - 29.00 89.88 5.5e+03 18124 ! U26736 Caenorhabditis eleg
gb_in1:DROMHC - 29.00 88.12 6.9e+03 22663 ! M61229 D.melanogaster myos
seq_name: gb_ov:AF012770

seq_documentation_block:

LOCUS AF012770 306 bp DNA VRT 25-OCT-1997
DEFINITION Channa gachua GAC4 cytochrome b (cytb) gene, mitochondrial
encoding mitochondrial protein, partial cds.

ACCESSION

AF012770.1 GI:2564448

VERSION

AF012770.1 GI:2564448

KEYWORDS

Channa gachua.

SOURCE

Mitochondrion Channa gachua

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;

Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Perciformes;

Perciformes; Channidae; Channidae; Channa.

REFERENCE

1 (bases 1 to 306)

AUTHORS

Chan, W.C. and Lee, P.G.

TITLE

Systematics of southeast Asian snakeheads using molecular and morphological data

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 306)

AUTHORS

Chan, W.C. and Lee, P.G.

TITLE

Direct Submission

JOURNAL

Submitted (08-JUL-1997) School of Biological Sciences, National University of Singapore, 10 Kent Ridge Crescent 119260, Singapore

FEATURES

Location/Qualifiers

1..306

/organism="Channa gachua"

/mitochondrion

/isolate="GAC4"

/db_xref="taxon:33790"

/note="from Nee Soon Swamp Forest, Singapore"

<1..>306

/gene="cytb"

<1..>306

/gene="cytb"

/codon_start=1

/translation="FGLSLGLCLTQLTGLFLAMHYTSDISTAFSSVAHICRDVNYG"

WLIIRNLHANGASFFIICIIYHIGRLYGVLYKWTWNGVVMVLLVMTAFVGVLP

..

BASE COUNT 73 a 80 c 53 g 100 t

ORIGIN

1

alignment_scores:

Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-6 x AF012770 ..

Align seg 1/1 to: AF012770 from: 1 to: 306

1 TyrArgLeuLeuIleArg 6

84 TACCGCCTTCATCCGT 101

seq_name: gb_ov:AF012771

seq_documentation_block:

LOCUS AF012771 306 bp DNA VRT 25-OCT-1997

DEFINITION Channa gachua GAC6 cytochrome b (cytb) gene, mitochondrial

encoding mitochondrial protein, partial cds.

ACCESSION

AF012771

VERSION

AF012771.1 GI:2564450

```

KEYWORDS      Channa gachua.
SOURCE        Mitochondrion Channa gachua
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
              Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
              Perciformes; Channoidei; Channidae; Channa.
REFERENCE     1 (bases 1 to 306)
AUTHORS       Chan,W.C. and Lee,P.G.
TITLE         Systematics of southeast Asian snakeheads using molecular and
              morphological data
JOURNAL       Unpublished
REFERENCE     2 (bases 1 to 306)
AUTHORS       Chan,W.C. and Lee,P.G.
TITLE         Direct Submission
JOURNAL       Submitted (08-JUL-1997) School of Biological Sciences, National
              University of Singapore, 10 Kent Ridge Crescent 119260, Singapore
FEATURES     Location/Qualifiers
              source
                1..306
                /organism="Channa gachua"
                /mitochondrion
                /isolate="GAC6"
                /db_xref="taxon:33790"
                /note="from Kuala Lumpur, Malaysia"
                <1..>306
                /gene="cytb"
                <1..>306
                /gene="cytb"
                /codon_start=1
                /transl_table=2
                /product="cytochrome b"
                /protein_id="AA81734.1"
                /db_xref="GI:2584451"
                /translation="FGLSLGLWLTQILTGFLAMHYTSDISTACSSVAHICRDVNYG
                WLIRNLHANGASFFICIFYFHIGRGLYGSILYKETWNVGVVLLVMTAFVGYVLP"
BASE COUNT   71 a 77 c 54 g 104 t
ORIGIN
alignment_scores:
  Quality: 29.00      Length: 6
  Ratio: 4.833       Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
  US-08-653-294-6 x AF012771 ..
  Align seg 1/1 to: AF012771 from: 1 to: 306
      1 TyArgLeuLeuIleArg 6
      |||||
      84 TACCGCTGCTCATCCGT 101
seq_name: gb_ov:AF012771
seq_documentation_block:
  LOCUS      AF012771      306 bp      DNA      VRT      25-OCT-1997
  DEFINITION Channa micropeltes Mic1 cytochrome b (cytb) gene, mitochondrial
              gene encoding mitochondrial protein, partial cds.
  ACCESSION  AF012771
  VERSION    AF012771.1 GI:2564476
  KEYWORDS
  SOURCE     Channa micropeltes.
  ORGANISM   Mitochondrion Channa micropeltes
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
              Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
              Perciformes; Channoidei; Channidae; Channa.
  REFERENCE  1 (bases 1 to 306)
  AUTHORS    Chan,W.C. and Lee,P.G.
  TITLE      Systematics of southeast Asian snakeheads using molecular and
              morphological data
  JOURNAL    Unpublished
  REFERENCE  2 (bases 1 to 306)
  AUTHORS    Chan,W.C. and Lee,P.G.
  TITLE      Direct Submission
  JOURNAL    Submitted (08-JUL-1997) School of Biological Sciences, National
              University of Singapore, 10 Kent Ridge Crescent 119260, Singapore
FEATURES     Location/Qualifiers
              source
                1..306
                /organism="Channa micropeltes"
                /mitochondrion
                /isolate="MIC1"
                /db_xref="taxon:64149"
                /note="from Singapore"
                <1..>306
                /gene="cytb"
                <1..>306
                /gene="cytb"
                /codon_start=1
                /transl_table=2
                /product="cytochrome b"
                /protein_id="AA81747.1"
                /db_xref="GI:2564477"
                /translation="FGLSLGLWLTQILTGFLAMHYTSDISTAFSSVAHICRDVNYG
                WLIRNLHANGASFFICIFYFHIGRGLYGSILYKETWNVGVVLLVMTACVGYVLP"
BASE COUNT   71 a 81 c 53 g 101 t
ORIGIN
alignment_scores:
  Quality: 29.00      Length: 6
  Ratio: 4.833       Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
  US-08-653-294-6 x AF012784 ..
  Align seg 1/1 to: AF012784 from: 1 to: 306
      1 TyArgLeuLeuIleArg 6
      |||||
      84 TACCGCTGCTCATCCGT 101
seq_name: gb_ov:AF012785
seq_documentation_block:
  LOCUS      AF012785      306 bp      DNA      VRT      25-OCT-1997
  DEFINITION Channa micropeltes MIC2 cytochrome b (cytb) gene, mitochondrial
              gene encoding mitochondrial protein, partial cds.
  ACCESSION  AF012785
  VERSION    AF012785.1 GI:2564478
  KEYWORDS
  SOURCE     Channa micropeltes.
  ORGANISM   Mitochondrion Channa micropeltes
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
              Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
              Perciformes; Channoidei; Channidae; Channa.
  REFERENCE  1 (bases 1 to 306)
  AUTHORS    Chan,W.C. and Lee,P.G.
  TITLE      Systematics of southeast Asian snakeheads using molecular and
              morphological data
  JOURNAL    Unpublished
  REFERENCE  2 (bases 1 to 306)
  AUTHORS    Chan,W.C. and Lee,P.G.
  TITLE      Direct Submission
  JOURNAL    Submitted (08-JUL-1997) School of Biological Sciences, National
              University of Singapore, 10 Kent Ridge Crescent 119260, Singapore
FEATURES     Location/Qualifiers
              source
                1..306
                /organism="Channa micropeltes"
                /mitochondrion
                /isolate="MIC2"
                /db_xref="taxon:64149"
                /note="from Singapore"
                <1..>306

```


CDS
/gene="cytb"
<1..>306
/gene="cytb"
/codon_start=1
/transl_table=2
/product="cytochrome b"
/protein_id="AAB81748.1"
/db_xref="GI:2564479"
/translation="FGSLGLCLITQILTGLFLAMHYTSDISTAFSSVAHICRDVNYG
WLIRNLHANGASFFICIFYFHGRGLYGSYLYKRTWNVGVMLLVMTACGVLP

BASE COUNT 74 a 76 c 54 g 102 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x AF012785 ..

Align seg 1/1 to: AF012785 from: 1 to: 306

1 TytArgLeuLeuIleArg 6
|||||
84 TACCGCTTCTCATCCGT 101

seq_name: gb_ov:AF012786

seq_documentation_block:
LOCUS AF012786 306 bp DNA VRT 25-OCT-1997
DEFINITION Channa orientalis ORL1 cytochrome b (cytb) gene, mitochondrial gene
encoding mitochondrial protein, partial cds.
ACCESSION AF012786
VERSION AF012786.1 GI:2564480
KEYWORDS
SOURCE Channa orientalis.
ORGANISM Mitochondrion Channa orientalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percormorpha;
Perciformes; Channidae; Channidae; Channa.

REFERENCE 1 (bases 1 to 306)
AUTHORS Chan,W.C. and Lee,P.G.
TITLE Systematics of Southeast Asian snakeheads using molecular and morphological data
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 306)
AUTHORS Chan,W.C. and Lee,P.G.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-1997) School of Biological Sciences, National University of Singapore, 10 Kent Ridge Crescent 119260, Singapore
FEATURES
source
1..306
/organism="Channa orientalis"
/mitochondrion
/isolate="ORL1"
/db_xref="taxon:64150"
/note="from Sri Lanka"
<1..>306
/gene="cytb"
<1..>306
/gene="cytb"
/codon_start=1
/transl_table=2
/product="cytochrome b"
/protein_id="AAB81749.1"
/db_xref="GI:2564481"
/translation="FGSLGLCLITQILTGLFLAMHYTSDISTAFSSVAHICRDVNYG
WLIRNLHANGASFFICIFYFHGRGLYGSYLYKRTWNVGVMLLVMTACGVLP

BASE COUNT 77 a 83 c 49 g 97 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x AF012787 ..

Align seg 1/1 to: AF012787 from: 1 to: 306

ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x AF012786 ..

Align seg 1/1 to: AF012786 from: 1 to: 306

1 TytArgLeuLeuIleArg 6
|||||
84 TACCGCTTCTCATCCGT 101

seq_name: gb_ov:AF012787

seq_documentation_block:
LOCUS AF012787 306 bp DNA VRT 25-OCT-1997
DEFINITION Channa orientalis OR12 cytochrome b (cytb) gene, mitochondrial gene
encoding mitochondrial protein, partial cds.
ACCESSION AF012787
VERSION AF012787.1 GI:2564482
KEYWORDS
SOURCE Channa orientalis.
ORGANISM Mitochondrion Channa orientalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percormorpha;
Perciformes; Channidae; Channidae; Channa.

REFERENCE 1 (bases 1 to 306)
AUTHORS Chan,W.C. and Lee,P.G.
TITLE Systematics of Southeast Asian snakeheads using molecular and morphological data
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 306)
AUTHORS Chan,W.C. and Lee,P.G.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-1997) School of Biological Sciences, National University of Singapore, 10 Kent Ridge Crescent 119260, Singapore
FEATURES
source
1..306
/organism="Channa orientalis"
/mitochondrion
/isolate="OR12"
/db_xref="taxon:64150"
/note="from Sri Lanka"
<1..>306
/gene="cytb"
<1..>306
/gene="cytb"
/codon_start=1
/transl_table=2
/product="cytochrome b"
/protein_id="AAB81750.1"
/db_xref="GI:2564483"
/translation="FGSLGLCLITQILTGLFLAMHYTSDISTAFSSVAHICRDVNYG
WLIRNLHANGASFFICIFYFHGRGLYGSYLYKRTWNVGVMLLVMTACGVLP

BASE COUNT 77 a 84 c 52 g 93 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x AF012787 ..

Align seg 1/1 to: AF012787 from: 1 to: 306

```

1 TyArgLeuLeuIleArg 6
|||||
84 TACCGCCTTCATCCGT 101

seq_name: gb_ov:AF012789

seq_documentation_block:
LOCUS AF012789 306 bp DNA VRT 25-OCT-1997
DEFINITION Channa striata STR1 cytochrome b (cytb) gene, mitochondrial gene
encoding mitochondrial protein, partial cds.
ACCESSION AF012789
VERSION AF012789.1 GI:2564486
KEYWORDS
SOURCE Channa striata.
ORGANISM Mitochondrion Channa striata
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percormorpha;
Perciformes; Channoidei; Channidae; Channa.
REFERENCE
1 (bases 1 to 306)
AUTHORS Chan,W.C. and Lee,P.G.
TITLE Systematics of southeast Asian snakeheads using molecular and
morphological data
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 306)
AUTHORS Chan,W.C. and Lee,P.G.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-1997) School of Biological Sciences, National
University of Singapore, 10 Kent Ridge Crescent 119260, Singapore
FEATURES
source
1..306
/organism="Channa striata"
/mitochondrion
/isolate="STRI"
/db_xref="taxon:64152"
/notes="from Air Hitam, Malaysia"
<1..>306
/gene="cytb"
<1..>306
/gene="cytb"
/codon_start=1
/transl_table=2
/product="cytochrome b"
/protein_id="AAB81752.1"
/db_xref="GI:2564487"
/translation="FSGSLGLCLITOLIGLFLAMHYTSDISTAFSSVAHICRDVNYG
WLIRNLHANGASFFICIFYFHIGRGLYGSILYKETWNVGMILLLVMTAFVGYLVP"

BASE COUNT 74 a 82 c 54 g 96 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x AF012789 ..
Align seg 1/1 to: AF012789 from: 1 to: 306

1 TyArgLeuLeuIleArg 6
|||||
84 TACCGCCTTCATCCGT 101

seq_name: gb_ov:CGMTCYB1

seq_documentation_block:
LOCUS CGMTCYB1 306 bp DNA VRT 24-FEB-1994
DEFINITION C.gachua (GAC/1/SING) mitochondrial gene for cytochrome b.
ACCESSION Z30267
VERSION Z30267.1 GI:456619
KEYWORDS
SOURCE Channa gachua.
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percormorpha;
Perciformes; Channoidei; Channidae; Channa.
REFERENCE
1 (bases 1 to 306)
AUTHORS Lee,P.G. and Chan,W.K.
TITLE Methods for obtaining cytochrome b gene sequences from fresh and
preserved tissues of snakehead fishes (Pisces; Channidae) using the
polymerase chain reaction (PCR)
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 306)
AUTHORS Chan,W.K.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-1994) W. K. Chan, Zoology, National University of
Singapore, Lower Kent, Ridge Road, Singapore, S0511, Republic Of
Singapore
FEATURES
source
1..306
/organism="Channa gachua"
/isolate="GAC/1/SING"
/db_xref="taxon:33790"
/dev_stage="adult"
/tissue_type="muscle, skeletal"
/clone_lib="PCR Fragment"

BASE COUNT 75 a 78 c 48 g 105 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x CGMTCYB1 ..
Align seg 1/1 to: CGMTCYB1 from: 1 to: 306

1 TyArgLeuLeuIleArg 6
|||||
84 TACCGCCTTCATCCGT 101

seq_name: gb_ov:CGMTCYB2

seq_documentation_block:
LOCUS CGMTCYB2 306 bp DNA VRT 24-FEB-1994
DEFINITION C.gachua (GAC/2/SRIL) mitochondrial gene for cytochrome b.
ACCESSION Z30268
VERSION Z30268.1 GI:456620
KEYWORDS
SOURCE Cytochrome b.
ORGANISM Channa gachua.
Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percormorpha;
Perciformes; Channoidei; Channidae; Channa.
REFERENCE
1 (bases 1 to 306)
AUTHORS Lee,P.G. and Chan,W.K.
TITLE Methods for obtaining cytochrome b gene sequences from fresh and
preserved tissues of snakehead fishes (Pisces; Channidae) using the
polymerase chain reaction (PCR)
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 306)
AUTHORS Chan,W.K.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-1994) W. K. Chan, Zoology, National University of
Singapore, Lower Kent, Ridge Road, Singapore, S0511, Republic Of
Singapore
FEATURES
source
1..306
/organism="Channa gachua"
/isolate="GAC/2/SRIL"

```

seq_name: gb_ov:CGMTCYB3
 seq_documentation_block:
 LOCUS CGMTCYB3 306 bp DNA VRT 24-FEB-1994
 DEFINITION C.gachua (GAC/3/SRIL) mitochondrial gene for cytochrome b.
 ACCESSION Z30269
 VERSION Z30269.1 GI:456621
 KEYWORDS cytochrome b.
 SOURCE Channa gachua.
 ORGANISM Channa gachua.
 Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
 Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percormorpha;
 Perciformes; Channoidel; Channidae; Channa.
 REFERENCE 1 (bases 1 to 306)
 AUTHORS Lee, P.G. and Chan, W.K.
 TITLE Methods for obtaining cytochrome b gene sequences from fresh and preserved tissues of snakehead fishes (Pisces; Channidae) using the polymerase chain reaction (PCR)

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-6 x CGMTCYB2

Align seg 1/1 to: CGMTCYB2 from: 1 to: 306

1 TyrArgLeuLeuIleArg 6
 |||||
 84 TACCGCTTCTCATCGT 101

seq_name: gb_ov:CGMTCYB3

seq_documentation_block:
 LOCUS CGMTCYB3 306 bp DNA VRT 24-FEB-1994
 DEFINITION C.gachua (GAC/3/SRIL) mitochondrial gene for cytochrome b.
 ACCESSION Z30269
 VERSION Z30269.1 GI:456621
 KEYWORDS cytochrome b.
 SOURCE Channa gachua.
 ORGANISM Channa gachua.
 Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
 Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percormorpha;
 Perciformes; Channoidel; Channidae; Channa.
 REFERENCE 1 (bases 1 to 306)
 AUTHORS Lee, P.G. and Chan, W.K.
 TITLE Methods for obtaining cytochrome b gene sequences from fresh and preserved tissues of snakehead fishes (Pisces; Channidae) using the polymerase chain reaction (PCR)

alignment_block:
 US-08-653-294-6 x CGMTCYB4

Align seg 1/1 to: CGMTCYB4 from: 1 to: 306

1 TyrArgLeuLeuIleArg 6
 |||||
 84 TACCGCTTCTCATCGT 101

FEATURES

source
 1..306
 /organism="Channa gachua"
 /isolate="GAC/3/SRIL"
 /db_xref="taxon:33790"
 /dev_stage="Adult"
 /tissue_type="Muscle, skeletal"
 /clone_lib="PCR Fragment"
 BASE COUNT 75 a 79 c 49 g 103 t
 ORIGIN

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-6 x CGMTCYB3

Align seg 1/1 to: CGMTCYB3 from: 1 to: 306

1 TyrArgLeuLeuIleArg 6
 |||||
 84 TACCGCTTCTCATCGT 101

seq_name: gb_ov:CGMTCYB4

seq_documentation_block:
 LOCUS CGMTCYB4 306 bp DNA VRT 24-FEB-1994
 DEFINITION C.gachua (GAC/5/KUAL) mitochondrial gene for cytochrome b.
 ACCESSION Z30270
 VERSION Z30270.1 GI:456644
 KEYWORDS cytochrome b.
 SOURCE Channa gachua.
 ORGANISM Channa gachua.
 Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
 Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percormorpha;
 Perciformes; Channoidel; Channidae; Channa.
 REFERENCE 1 (bases 1 to 306)
 AUTHORS Lee, P.G. and Chan, W.K.
 TITLE Methods for obtaining cytochrome b gene sequences from fresh and preserved tissues of snakehead fishes (Pisces; Channidae) using the polymerase chain reaction (PCR)

alignment_block:

US-08-653-294-6 x CGMTCYB4

Align seg 1/1 to: CGMTCYB4 from: 1 to: 306

1 TyrArgLeuLeuIleArg 6
 |||||
 84 TACCGCTTCTCATCGT 101

FEATURES

source
 1..306
 /organism="Channa gachua"
 /isolate="GAC/5/KUAL"
 /db_xref="taxon:33790"
 /dev_stage="adult"
 /tissue_type="muscle, skeletal"
 /clone_lib="PCR Fragment"
 BASE COUNT 72 a 77 c 52 g 105 t
 ORIGIN

alignment_scores:

Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-6 x CGMTCYB4

Align seg 1/1 to: CGMTCYB4 from: 1 to: 306

1 TyrArgLeuLeuIleArg 6
 |||||
 84 TACCGCTTCTCATCGT 101

seq_name: gb_ov:CGMTCYB5

seq_documentation_block:
 LOCUS CGMTCYB5 306 bp DNA VRT 24-FEB-1994
 DEFINITION C.gachua (GAC/7/SING) mitochondrial gene for cytochrome b.
 ACCESSION Z30271
 VERSION Z30271.1 GI:456645
 KEYWORDS cytochrome b.
 SOURCE Channa gachua.
 ORGANISM Channa gachua.
 Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
 Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percormorpha;
 Perciformes; Channoidel; Channidae; Channa.
 REFERENCE 1 (bases 1 to 306)
 AUTHORS Lee, P.G. and Chan, W.K.
 TITLE Methods for obtaining cytochrome b gene sequences from fresh and preserved tissues of snakehead fishes (Pisces; Channidae) using the polymerase chain reaction (PCR)

alignment_block:

US-08-653-294-6 x CGMTCYB5

Align seg 1/1 to: CGMTCYB5 from: 1 to: 306

1 TyrArgLeuLeuIleArg 6
 |||||
 84 TACCGCTTCTCATCGT 101

FEATURES

source

JOURNAL Submitted (21-FEB-1994) W. K. Chan, Zoology, National University of Singapore, Lower Kent, Ridge Road, Singapore, S0511, Republic Of Singapore

FEATURES

source Location/Qualifiers
 1..306
 /organism="Channa gachua"
 /isolate="GAC/7/SING"
 /db_xref="taxon:33790"
 /dev_stage="adult"
 /tissue_type="muscle, skeletal"
 /clone_lib="PCR Fragment"
 BASE COUNT 74 a 79 c 50 g 103 t
 ORIGIN

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-6 x CGMTCYB5 ..

Align seg 1/1 to: CGMTCYB5 from: 1 to: 306

1 TyrArgLeuLeulleArg 6

|||||
 84 TACCGCTGCTCATCCGT 101

seq_name: gb_ov:PAU26956

seq_documentation_block: 340 bp DNA VRT 23-JUL-1995
 LOCUS PAU26956
 DEFINITION Pristipomoides aquilonaris cytochrome b gene, mitochondrial gene encoding mitochondrial protein, partial cds.

ACCESSION U26956
 VERSION U26956.1 GI:903362

KEYWORDS
 SOURCE

ORGANISM

Pristipomoides aquilonaris.
 Mitochondrion Pristipomoides aquilonaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
 Perciformes; Percoidae; Lutjanidae; Pristipomoides.

1 (bases 1 to 340)

Salver, S.K., Freshwater, D. and Walsh, P.J.

Phylogenetic relationships of western Atlantic snappers (Family

Lutjanidae) based on mitochondrial DNA sequences

Unpublished

2 (bases 1 to 340)

Freshwater, D.W.

Direct Submission

Submitted (12-MAY-1995) David W. Freshwater, Center for Marine

Science Research, 7205 Wrightsville Avenue, Wilmington, NC 28403,

USA

FEATURES Location/Qualifiers

1..340

/organism="Pristipomoides aquilonaris"

/mitochondrion

/db_xref="taxon:40511"

<1..>340

/codon_start=3

/transl_table=2

/product="cytochrome b"

/protein_id="AAA0184.1"

/db_xref="GI:903363"

/translation="SAWNFGSLGLCLAAQILTLFLAMHYTSDISAFSYVAHICR

DVNYGLIRNLHANGASFFFCILYHIGRGLYGSYLKYNIGVLELLVMTAFV

GYVLPXGQMS"

BASE COUNT 77 a 104 c 56 g 102 t 1 others

ORIGIN

alignment_scores:

Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-6 x PAU26956 ..

Align seg 1/1 to: PAU26956 from: 1 to: 340

1 TyrArgLeuLeulleArg 6

|||||
 101 TATCGCTTCTCATCGT 118

seq_name: gb_ov:FDU77119

seq_documentation_block: 357 bp DNA VRT 07-APR-1998

LOCUS FDU77119

DEFINITION Fundulus dispar cytochrome b (cytb) gene, mitochondrial gene encoding mitochondrial protein, partial cds.

ACCESSION U77119

VERSION U77119.1 GI:2329956

KEYWORDS

SOURCE

ORGANISM

Fundulus dispar.

Mitochondrion Fundulus dispar

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;

Neopterygii; Teleostei; Euteleostei; Acanthopterygii;

Atherinomorpha; Cyprinodontiformes; Cyprinodontidae; Fundulidae;

Fundulus.

1 (bases 1 to 357)

Ghedotti, M.J. and Grose, M.J.

Phylogenetic relationships of the Fundulus notii species group

(Fundulidae, Cyprinodontiformes) as inferred from the cytochrome-b

gene

2 (bases 1 to 357)

Ghedotti, M.J. and Grose, M.J.

Direct Submission

Submitted (04-NOV-1996) Natural History Museum, University of

Kansas, 502E Dyche Hall, Lawrence, KS 66045-2454, USA

FEATURES Location/Qualifiers

1..357

/organism="Fundulus dispar"

/mitochondrion

/db_xref="taxon:34778"

/tissue_type="muscle"

/note="individual 1"

1..357

/gene="cytb"

1..>357

/gene="cytb"

/codon_start=1

/transl_table=2

/product="cytochrome b"

/protein_id="AAC12689.1"

/db_xref="GI:2329957"

/translation="MANIRKTHPLFKIVNNALVDLPAPVNISVWMNFGSLGLCLMSQ

ILTGLFLAMHYTSDISTAFSSVWHICRDVNTGWLRNLHANGASFFFCIVLHIGRGL

YGSYLKYNIGVLELLVMTAFV"

BASE COUNT 96 a 77 c 56 g 128 t

ORIGIN

alignment_scores:

Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-6 x FDU77119 ..

Align seg 1/1 to: FDU77119 from: 1 to: 357

1 TyrArgLeuLeulleArg 6

```
|||||
180 TACCGCTTCTCATCGT 197

seq_name: gb_ov:FDU77120

seq_documentation_block: 357 bp DNA VRT 07-APR-1998
LOCUS FDU77120
DEFINITION Fundulus dispar cytochrome b (cytb) gene, mitochondrial gene
encoding mitochondrial protein, partial cds.
ACCESSION U77120
VERSION U77120.1 GI:2329958
KEYWORDS
SOURCE
ORGANISM
Fundulus dispar.
Mitochondrion Fundulus dispar
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Fundulidae;
Atherinomorphi; Cyprinodontiformes; Cyprinodontoidae; Fundulidae;
Fundulus.
REFERENCE
1 (bases 1 to 357)
Ghedotti, M.J. and Grose, M.J.
Phylogenetic relationships of the Fundulus nottii species group
(Fundulidae, Cyprinodontiformes) as inferred from the cytochrome-b
gene
Copeia 4, 858-862 (1997)
REFERENCE
2 (bases 1 to 357)
Ghedotti, M.J. and Grose, M.J.
Direct Submission
TITLE
Submitted (04-NOV-1996) Natural History Museum, University of
Kansas, 502E Dyche Hall, Lawrence, KS 66045-2454, USA
LOCATION/Qualifiers
1. .357
/organism="Fundulus dispar"
/mitochondrion
/db_xref="taxon:34778"
/tissue_type="muscle"
/note="individual 2"
1. .357
/gene="cytb"
1. .357
/gene="cytb"
/codon_start=1
/transl_table=2
/product="cytochrome b"
/protein_id="AAC12690.1"
/db_xref="GI:2329958"
/translation="MANIKRTHPLKIVNNALVDLPAPVNISSWNFGSLGLCLISQ
ILTGFLAHYTSIDISTAFSSVHICRDVNYGWLIRNHANGASEFFICIIYHIGRGL
YYGSYLYKRETNVGVIL"

BASE COUNT 96 a 77 c 55 g 129 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x FDU77120 ..
Align seg 1/1 to: FDU77120 from: 1 to: 357
1 TyrArgLeuLeuIleArg 6
|||||
180 TACCGCTTCTCATCGT 197
```

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-6 to: N_Geneseq_36:* out_format : pfs
Date: Feb 8, 2000 1:27 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-Q/cgml1/USPTO/spool/US08653294/runat_04022000_160701_15807/app_query.fasta.1
-DB-N_Geneseq_36 -QMT-fastcap -SUFFIX-rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-GAPOP=6.000 -GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX-blosum62
-TRANS-human40.coi -LIST=45 -DALIGN=200 -THR_SCORE=epct
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=100000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT
-THREADS=1

Search information block:
Query: US-08-653-294-6
Query length: 6
Database: N_Geneseq_36:*
Database sequences: 311585
Database length: 125096042
Search time (sec): 590.520000

| score_list: | Sequence | Strd | Orig | ZScore | Escore | Len | Documentation |
|-------------|----------------------|------|-------|--------|---------|-------|-----------------------------------|
| | N_Geneseq_36:V86097 | - | 29.00 | 112.40 | 116.30 | 506 | EST clone D310. New polynucleot |
| | N_Geneseq_36:X20560 | + | 29.00 | 88.92 | 2.4e+03 | 8467 | Polynucleotide sequence from t |
| | N_Geneseq_36:V31189 | - | 29.00 | 88.65 | 2.4e+03 | 8752 | E. coli J96 pathogenicity isla |
| | N_Geneseq_36:Q54386 | + | 29.00 | 74.66 | 1.5e+04 | 46899 | T. niveum Cyclosporin synthet |
| | N_Geneseq_36:R22714 | - | 28.00 | 115.33 | 79.88 | 215 | Human gene signature HUMG504372 |
| | N_Geneseq_36:R43366 | - | 28.00 | 107.99 | 204.90 | 519 | Cotton FbLate 2-82A gene cDNA c |
| | N_Geneseq_36:R75248 | - | 28.00 | 106.57 | 245.65 | 615 | Staphylococcus aureus contig SE |
| | N_Geneseq_36:R84061 | + | 28.00 | 105.69 | 275.22 | 684 | DNA encoding a Staphylococcus a |
| | N_Geneseq_36:R543095 | + | 28.00 | 102.26 | 427.18 | 1032 | DNA encoding a Staphylococcus p |
| | N_Geneseq_36:R13048 | - | 28.00 | 100.45 | 539.10 | 1283 | Cotton fibre-specific cDNA c |
| | N_Geneseq_36:R30265 | - | 28.00 | 100.45 | 539.10 | 1283 | Cotton fibre cell-specific cDN |
| | N_Geneseq_36:R62624 | - | 28.00 | 100.45 | 539.10 | 1283 | Cotton fibre specific cDNA c |
| | N_Geneseq_36:R70055 | - | 28.00 | 100.45 | 539.10 | 1283 | Cotton fibre specific cDNA c |
| | N_Geneseq_36:V44448 | - | 28.00 | 95.34 | 1.0e+03 | 2367 | Mycobacterium tuberculosis ant |
| | N_Geneseq_36:V54236 | + | 28.00 | 95.34 | 1.0e+03 | 2367 | M. tuberculosis immunogenic p |
| | N_Geneseq_36:V52296 | + | 28.00 | 84.69 | 4.1e+03 | 8494 | Streptococcus pneumoniae genom |
| | N_Geneseq_36:V15570 | + | 27.00 | 125.57 | 21.48 | 38 | Primer for HIV RNA. Inhibiting r |
| | N_Geneseq_36:Q27321 | + | 27.00 | 124.16 | 25.74 | 45 | T cell receptor beta gene rearra |
| | N_Geneseq_36:V64315 | + | 27.00 | 124.16 | 25.74 | 45 | Human TCR Beta chain rearranged |
| | N_Geneseq_36:R87598 | + | 27.00 | 107.43 | 220.03 | 335 | EST clone DY988. New polynucleo |
| | N_Geneseq_36:R75035 | + | 27.00 | 104.72 | 311.67 | 454 | Staphylococcus aureus contig SE |
| | N_Geneseq_36:R01698 | + | 27.00 | 102.53 | 412.41 | 603 | Human anti-angiogenic hPRL Met- |
| | N_Geneseq_36:X01701 | + | 27.00 | 102.53 | 412.41 | 603 | Human anti-angiogenic 16K hPRL |
| | N_Geneseq_36:R39936 | - | 27.00 | 101.01 | 501.45 | 724 | Gastric cancer associated gene. |
| | N_Geneseq_36:V59910 | + | 27.00 | 100.53 | 533.35 | 767 | DNA encoding Staphylococcus au |
| | N_Geneseq_36:V58934 | + | 27.00 | 99.48 | 610.24 | 870 | Humanised A33 scab 233 DNA. New |
| | N_Geneseq_36:R83894 | + | 27.00 | 99.06 | 644.04 | 915 | DNA encoding a Staphylococcus a |
| | N_Geneseq_36:R82078 | + | 27.00 | 96.13 | 937.43 | 1300 | DNA encoding a partial gldAI p |
| | N_Geneseq_36:X07180 | + | 27.00 | 95.06 | 1.1e+03 | 1478 | Corn threonine deaminase cDNA |
| | N_Geneseq_36:R794525 | - | 27.00 | 94.87 | 1.1e+03 | 1512 | Borna disease virus (BDV) prot |
| | N_Geneseq_36:V58937 | - | 27.00 | 94.34 | 1.2e+03 | 1611 | Z39523237Az chimeric receptor |
| | N_Geneseq_36:R69235 | - | 27.00 | 93.82 | 1.3e+03 | 1716 | Aspartase-470 coding sequence. |
| | N_Geneseq_36:Q14978 | + | 27.00 | 93.37 | 1.3e+03 | 1810 | Acylase gene. Gene encoding 3-acy |
| | N_Geneseq_36:R44560 | + | 27.00 | 93.12 | 1.4e+03 | 1842 | Human secreted protein gene 15 |
| | N_Geneseq_36:V23319 | + | 27.00 | 93.23 | 1.4e+03 | 1866 | Helicobacter pylori dapE gene |
| | N_Geneseq_36:R82077 | + | 27.00 | 92.01 | 1.6e+03 | 2132 | DNA encoding a gldAI protein. |
| | N_Geneseq_36:R311994 | + | 27.00 | 91.40 | 1.7e+03 | 2295 | Nonsense-mediated mRNA decay 2 |
| | N_Geneseq_36:R17044 | + | 27.00 | 90.91 | 1.8e+03 | 2434 | Bacillus thuringiensis insecti |
| | N_Geneseq_36:R74680 | + | 27.00 | 90.62 | 1.9e+03 | 2518 | Staphylococcus aureus contig S |
| | N_Geneseq_36:R38106 | - | 27.00 | 90.17 | 2.0e+03 | 2658 | Borna disease virus HE/80 part |
| | N_Geneseq_36:V74488 | + | 27.00 | 89.62 | 2.2e+03 | 2841 | Staphylococcus aureus contig 5 |

| | | | | | | |
|---|---|---------------|-------------------|---------|-------|-----------------------------|
| N_Geneseq_36:T08745 | - | 27.00 | 89.53 | 2.2e+03 | 2872 | Polycystronic maize UI4.1 D |
| N_Geneseq_36:T14849 | - | 27.00 | 89.53 | 2.2e+03 | 2872 | Polycystronic maize UI4.1 D |
| N_Geneseq_36:Q05328 | + | 27.00 | 87.83 | 2.7e+03 | 3519 | Rat tumour necrosis factor |
| N_Geneseq_36:Q87444 | + | 27.00 | 87.74 | 2.8e+03 | 3560 | Drosophila semaphorin I CDN |
| seq_name: N_Geneseq_36:V86097 | | | | | | |
| seq_documentation_block: | | | | | | |
| ID | V86097 | standard: | cDNA; | 506 | BP. | |
| AC | V86097; | | | | | |
| DT | 27-APR-1999 | (first entry) | | | | |
| DE | EST clone D310. | | | | | |
| KW | Expressed sequence tag; secreted protein; haematopoiesis regulator; | | | | | |
| KW | Tissue growth; activin; inhibin; tumour invasion suppressor; EST; human; | | | | | |
| KW | chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis; | | | | | |
| KW | receptor; ligand; anti-inflammatory; tumour inhibitor; ds. | | | | | |
| OS | Homo sapiens. | | | | | |
| PN | W09845435-A2. | | | | | |
| PD | 15-OCT-1998. | | | | | |
| PF | 10-APR-1998; | U06954. | | | | |
| PR | 10-APR-1997; | US-835913. | | | | |
| PA | (GEMV) GENETICS INST INC. | | | | | |
| PI | Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D, | | | | | |
| PI | Racie LA, Spaulding V, Treacy M; | | | | | |
| DR | WPI; 99-070076/06. | | | | | |
| PT | New polynucleotides encoding human secreted proteins - derived from | | | | | |
| PT | e.g. human blood, kidney, foetal lung, placenta, testes, brain, | | | | | |
| PS | Claim 1; Page 119; 633pp; English. | | | | | |
| CC | This sequence represents an expressed sequence tag (EST), and is a | | | | | |
| CC | polynucleotide of the invention. The polynucleotides of the invention are | | | | | |
| CC | all secreted EST sequences isolated from a variety of human tissue | | | | | |
| CC | sources. The EST sequences and proteins encoded by them are predicted to | | | | | |
| CC | have useful biological activities which would make them suitable for | | | | | |
| CC | treatment, preventing or ameliorating medical conditions in humans and | | | | | |
| CC | animals, although no supporting data is given. Suggested activities | | | | | |
| CC | include nutritional activity, immune stimulating or suppressing activity, | | | | | |
| CC | haematopoiesis regulating activity, tissue growth activity, | | | | | |
| CC | activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic | | | | | |
| CC | and thrombolytic activity, receptor/ligand activity, anti-inflammatory | | | | | |
| CC | activity, cadherin/tumour invasion suppressor activity, tumour inhibition | | | | | |
| CC | activity. The EST sequences are also stated to be useful for gene | | | | | |
| CC | therapy. | | | | | |
| SQ | Sequence | 506 BP; | 135 A; | 133 C; | 88 G; | 136 T; |
| alignment_scores: | | | | | | |
| | Quality: | 29.00 | Length: | 6 | | |
| | Ratio: | 4.833 | Gaps: | 0 | | |
| | Percent Similarity: | 100.000 | Percent Identity: | 100.000 | | |
| alignment_block: | | | | | | |
| | US-08-653-294-6 x V86097/rev | .. | | | | |
| Align seg 1/1 to reverse of: V86097 from: 1 to: 506 | | | | | | |
| | 1 TyrArgLeuLeuLeuArg 6 | | | | | |
| | | | | | | |
| | 133 TACCGTTGCTCATTTAGA 116 | | | | | |
| seq_name: N_Geneseq_36:X20560 | | | | | | |
| seq_documentation_block: | | | | | | |
| ID | X20560 | standard: | DNA; | 8467 | BP. | |
| AC | X20560; | | | | | |
| DT | 05-MAY-1999 | (first entry) | | | | |
| DE | Polynucleotide sequence from the genome of Treponema pallidum. | | | | | |
| KW | Treponema pallidum infection; syphilis; Borrelia infection; animal; | | | | | |
| KW | enzyme production; ds. | | | | | |
| KW | Treponema pallidum. | | | | | |
| PN | W09859034-A2. | | | | | |
| PD | 30-DEC-1998. | | | | | |
| PF | 23-JUN-1998; | UI3041. | | | | |

```

PR 24-JUN-1997; US-050667.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Fraser CH;
DR WPI: 99-081273/07.
PT New isolated Treponema pallidum nucleic acids - used to develop
PT products for the detection, diagnosis, characterisation, prevention
PT and therapy of T. pallidum infections, particularly syphilis
PS Claim 1: Page 523-527; 1150pp; English.
CC X20500-21243 represent polynucleotide sequences from the genome of
CC Treponema pallidum. The sequences can be used for detection,
CC diagnosis, characterisation, prevention and therapy for T. pallidum
CC infections, particularly syphilis. They can also be used for detecting
CC diseases related to Borrelia infections in animals, and for the
CC production of biosynthetic products such as enzymes.
SQ Sequence 8467 BP; 1979 A; 2308 C; 2087 G; 2081 T;

alignment_scores:
  Quality: 29.00 Length: 6
  Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x X20560
Align seg 1/1 to: X20560 from: 1 to: 8467

1 TyArgLeuLeuIleArg 6
|||||
7979 TATCGCTTTTGATCGG 7996

seq_name: N_Geneseq_36:V31189
seq_documentation_block:
ID V31189 standard; DNA; 8752 BP.
AC V31189;
DT 01-OCT-1998 (first entry)
DE E. coli J96 pathogenicity island contig #3.
KW PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pHER;
KW PAI V; phev; vaccine; protective immune response; ds.
OS Escherichia coli.
PN WO9822575-A2.
PF 21-NOV-1997; U21347.
PR 14-OCT-1997; US-061953.
PR 22-NOV-1996; US-031626.
PA (UYWI-) UNIV WISCONSIN.
PI Choi GH, Dillon PJ, Welch RA;
DR WPI: 98-312461/27.
PT New isolated uropathogenic E. coli nucleotide sequences - used to
PT develop products for the detection of pathogenic E. coli and to
PT elicit an immune response to pathogenic E. coli
PS Claim 21; Page 84-88; 250pp; English.
CC This sequence represents a E. coli strain J96 contig containing
CC pathogenicity island (PAI) sequences, and represents a nucleic acid
CC molecule of the invention. PAIs are large fragments of DNA which comprise
CC pathogenicity determinants. The sequences of the invention are taken from
CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near phev)
CC on the E. coli chromosome and is greater than 170 kb. PAI V is located at
CC approximately 94 min (at pHER) on the E. coli chromosome and is
CC approximately 160 kb in size. Antibodies specific to the proteins encoded
CC by the PAI open reading frames of the invention can be used in kits to
CC detect uropathogenic E. coli. The proteins are used in vaccines to elicit
CC a protective immune response in an animal to the uropathogenic E. coli
CC strain J96.
SQ Sequence 8752 BP; 2212 A; 2222 C; 2071 G; 2229 T;

alignment_scores:
  Quality: 29.00 Length: 6
  Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
US-08-653-294-6 x V31189/rev
Align seg 1/1 to reverse of: V31189 from: 1 to: 8752

1 TyArgLeuLeuIleArg 6
|||||
8544 TACGACTTTTGATACGT 8527

seq_name: N_Geneseq_36:Q54386
seq_documentation_block:
ID Q54386 standard; DNA; 46899 BP.
AC Q54386;
DT 08-JUL-1994 (first entry)
DE T. niveum Cyclosporin synthetase gene.
KW Enzyme; cyclosporin; synthetase-like activity; Tolypocladium niveum;
KW T. inflatum GAMS; biosynthesis; vector; cyclosporin synthetase; ss.
OS Tolypocladium niveum.
FH Key Location/Qualifiers
FT cds 885..46730
FT /*tag= a
FT /product= Cyclosporin synthetase
FT 40239..43129
FT /*tag= b
FT /note= "SalI restriction fragment, preferred
FT fragment, Claim 4"
FT misc_feature 37781..40244
FT /*tag= d
FT /note= "SalI restriction fragment, preferred
FT fragment, Claim 5"
FT EP-578616-A.
PN 12-JAN-1994.
PD 05-JUL-1993; 810474.
PR 09-JUL-1992; AT-001403.
PR 08-MAR-1993; AT-000437.
PR 29-APR-1993; CH-001310.
PR 04-MAY-1993; CH-001375.
PA (SANO ) SANDOZ LTD.
PA (SANO ) SANDOZ PATENT GMBH.
PA (SANO ) SANDOZ-ERFINDUNGEN VERM GES MBH.
PI Leitner E, Schneider E, Schoergendorfer K, Weber G;
DR WPI: 94-010432/02.
DR P-PSDB: R44929.
PT Isolated DNA sequence - which codes for enzyme having cyclosporin
PT synthetase like activity
PS Claim 6; Page 17-41; 93pp; English.
CC This sequence encodes an enzyme which has cyclosporin synthetase-
CC like activity. This sequence was isolated from Tolypocladium niveum
CC (formerly known as T. inflatum GAMS). The enzyme encoded by this
CC sequence catalyses the peptide biosynthesis of cyclosporins and
CC structurally related molecules. This sequence may be used for the
CC production of cyclosporin by transforming a vector containing this
CC sequence in to a recombinant host. This allows effective production
CC of antibiotic cyclosporin or its derivatives.
SQ Sequence 46899 BP; 10651 A; 13513 C; 12509 G; 10226 T;

alignment_scores:
  Quality: 29.00 Length: 6
  Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x Q54386
Align seg 1/1 to: Q54386 from: 1 to: 46899

1 TyArgLeuLeuIleArg 6
|||||
8698 TACCGACTTCTTTATACGG 8715

```



```

seq_name: N_Geneseq_36:T22714
seq_documentation_block:
ID T22714 standard: cDNA to mRNA; 215 BP.
AC T22714;
DE Human gene signature HUMGS04372.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-AL.
PD 01-JUN-1995.
PR 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1197; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 215 BP; 80 A; 23 C; 43 G; 67 T;

alignment_scores:
Quality: 28.00 Length: 6
Ratio: 4.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
US-08-653-294-6 x T22714/rev ..
Align seg 1/1 to reverse of: T22714 from: 1 to: 215

1 TyrArgLeuLeuArg 6
|||||
179 TACAGATTACTNGTCAGA 162

seq_name: N_Geneseq_36:T43366
seq_documentation_block:
ID T43366 standard: DNA; 519 BP.
AC T43366;
DE Cotton FbLate 2-82A gene cDNA clone All amplified fragment.
KW FbLate; promoter; fibre; transgenic plant; cotton;
KW Gossypium hirsutum; ds.
OS Synthetic.
PN WO9639021-A1.
PD 12-DEC-1996.
PF 06-JUN-1996; U09449.
PR 06-JUN-1995; US-467504.
PA (MONS) MONSANTO CO.
PI John ME;
WPI; 97-042726/04.

PT Plant fibre-specific, developmentally regulated FbLate promoter -
PT useful for producing transgenic plants, esp. cotton, with altered
PT fibre properties
PS Example 5; Page 63; 79pp; English.
CC A DNA clone (T43366) was generated by 5'RACE using primers (see
CC also T43364-65) based on FbLate2 clone All (T43362), a partial
CC cDNA clone corresponding to mRNA prevalent in the late development
CC of cotton fibre. The RACE product showed 91.6% similarity at the
CC nucleotide level to the genomic clone, FbLate2-82A (see also
CC T43360). The homology of the RACE product started from nucleotide
CC position 2269 of the FbLate2-28A gene. The ARG initiation codon
CC was identified at position 2315 of the gene.
SQ Sequence 519 BP; 191 A; 127 C; 87 G; 114 T;

alignment_scores:
Quality: 28.00 Length: 6
Ratio: 4.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
US-08-653-294-6 x T43366/rev ..
Align seg 1/1 to reverse of: T43366 from: 1 to: 519

1 TyrArgLeuLeuArg 6
|||||
177 TACCGATTATTAGTGAGA 160

seq_name: N_Geneseq_36:V75248
seq_documentation_block:
ID V75248 standard: DNA; 615 BP.
AC V75248;
DE 16-MAR-1999 (first entry)
DE Staphylococcus aureus contig SEQ ID #937.
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
OS Staphylococcus aureus.
FH Key Location/Qualifiers
FT misc_feature 361..420
FT /tag= a
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT
FT EP-786519-A2.
PD 30-JUL-1997.
PF 07-JAN-1997; 100117.
PR 03-JAN-1996; US-009861.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
PI Rosen CA;
WPI; 97-374922/35.
DR Polynucleotide(s) and proteins derived from Staphylococcus aureus
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
PS Claim 1; Page 1755-1756; 3271pp; English.
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock

```

CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S.aureus DNA sequences contained on the
 CC computer readable medium.
 SQ Sequence 615 BP; 202 A; 102 C; 83 G; 163 T;

alignment_scores:
 Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:

US-08-653-294-6 x V75248/rev ..

Align seg 1/1 to reverse of: V75248 from: 1 to: 615

```

1 TyrArgLeuLeuIleArg 6
|||||
122 TATAGATTGCTGTTCCG 105

```

seq_name: N_Geneseq_36:T84061

seq_documentation_block:

ID T84061 standard; DNA; 684 BP.

AC T84061.

DE 27-AUG-1998 (first entry)

DT DNA encoding a Staphylococcus aureus protein of unknown function.

KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;

KW Staphylococcal gene; regulatory element; bacterial gene expression;

KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;

KW toxic shock syndrome; ss.

OS Staphylococcus aureus.

FH Key Location/Qualifiers

FT CDS 125..436

FT /*tag= a

PN WO9730070-A1.

PD 21-AUG-1997.

PF 19-FEB-1997; U02318.

PR 20-FEB-1996; US-011888.

PA (SMIK) SMITHKLINE BEECHAM CORP.

PI Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO,

PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;

DR WPI: 97-424969/39.

DR P-PSDB; W8135.

PT Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used

PT to isolate antimicrobial compounds, and in vaccines against S.

PT aureus infection

PS Claim 9; Page 853; 989pp; English.

CC The present sequence encodes a Staphylococcus aureus protein of

CC unknown function. The present sequence was isolated from a

CC library of clones of S. aureus WCUH 29 in Escherichia coli. The DNA

CC sequence can be used in the construction of ribozymes and antisense

CC sequences to control the expression of Staphylococcal genes. The DNA

CC sequence is also useful as a source of regulatory elements for the

CC control of bacterial gene expression. The encoded protein may be used

CC to produce vaccines to enable a host to produce specific antibodies

CC with antibacterial action. These vaccines and antibodies would protect

CC a host against invasion by S. aureus, and conditions relating to

CC Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled

CC skin syndrome, and toxic shock syndrome.

SQ Sequence 684 BP; 211 A; 123 C; 118 G; 228 T;

alignment_scores:

Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:

US-08-653-294-6 x T84061/rev ..

Align seg 1/1 to reverse of: T84061 from: 1 to: 684

```

1 TyrArgLeuLeuIleArg 6
|||||
567 TACCCGCTGTTAGTCGC 550

```

seq_name: N_Geneseq_36:T84095

seq_documentation_block:

ID T84095 standard; DNA; 1032 BP.

AC T84095.

DE 27-AUG-1998 (first entry)

DT DNA encoding a Staphylococcus aureus protein of unknown function.

KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;

KW Staphylococcal gene; regulatory element; bacterial gene expression;

KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;

KW toxic shock syndrome; ss.

OS Staphylococcus aureus.

FH Key Location/Qualifiers

FT CDS 253..537

FT /*tag= a

PN WO9730070-A1.

PD 21-AUG-1997.

PF 19-FEB-1997; U02318.

PR 20-FEB-1996; US-011888.

PA (SMIK) SMITHKLINE BEECHAM CORP.

PI Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO,

PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;

DR WPI: 97-424969/39.

DR P-PSDB; W8135.

PT Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used

PT to isolate antimicrobial compounds, and in vaccines against S.

PT aureus infection

PS Claim 9; Page 874; 989pp; English.

CC The present sequence encodes a Staphylococcus aureus protein of

CC unknown function. The present sequence was isolated from a

CC library of clones of S. aureus WCUH 29 in Escherichia coli. The DNA

CC sequence can be used in the construction of ribozymes and antisense

CC sequences to control the expression of Staphylococcal genes. The DNA

CC sequence is also useful as a source of regulatory elements for the

CC control of bacterial gene expression. The encoded protein may be used

CC to produce vaccines to enable a host to produce specific antibodies

CC with antibacterial action. These vaccines and antibodies would protect

CC a host against invasion by S. aureus, and conditions relating to

CC Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled

CC skin syndrome, and toxic shock syndrome.

SQ Sequence 1032 BP; 349 A; 162 C; 182 G; 313 T;

alignment_scores:

Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:

US-08-653-294-6 x T84095 ..

Align seg 1/1 to: T84095 from: 1 to: 1032

```

1 TyrArgLeuLeuIleArg 6
|||||
655 TATAGACTATTAGTCGA 672

```

seq_name: N_Geneseq_36:T13048

seq_documentation_block:

ID T13048 standard; cDNA; 1283 BP.

AC T13048.

DE 27-MAY-1996 (first entry)

DE Cotton fibre-specific cDNA clone E9.

KW Cotton; fibre; promoter; transgenic plant; crop improvement; ds.

OS Gossypium hirsutum strain Coker 312.

PN US9495070-A.

PD 27-FEB-1996. 253243.
 PF 04-OCT-1988; US-253243.
 PR 04-OCT-1988; US-253243.
 PR 21-NOV-1990; US-617239.
 PR 18-MAY-1992; US-885970.
 PA (CETU) AGRACETUS INC.
 PI John M;
 DR WPI: 96-139095/14.
 PT New isolated fibre-specific promoters - used for introducing
 PT altered fibre-specific characteristics into plants, partic. cotton.
 PS Example 3; Column 45-46; 48pp; English.
 CC Cotton cDNA clone E9 (T13048) was isolated from a cDNA library of
 CC cotton var. Coker 312 15-day-old boll cells using a subtractive
 CC hybridization procedure. The clone hybridises strongly to fiber
 CC RNA and weakly to petal RA. E9 and other fibre-specific cDNA clones
 CC (see T13033-47 and T13049-T13050) were used to screen cotton genomic
 CC libraries, leading to the isolation of genomic clones (see T13025-32
 CC and T13052-53) contg. sequences capable of promoting gene expression
 CC in fibre cells.
 SQ Sequence 1283 BP; 509 A; 233 C; 251 G; 290 T;

alignment_scores:
 Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
 US-08-653-294-6 x T13048/rev ..

Align seg 1/1 to reverse of: T13048 from: 1 to: 1283

1 TyrArgLeuLeuIleArg 6
 |||||
 118 TACCGATTATTAGTGAGA 101

seq_name: N_Geneseq_36:T30265

seq_documentation_block:
 ID T30265 standard; DNA; 1283 BP.
 AC T30265;
 DT 16-DEC-1996 (first entry)
 DE Cotton fibre cell-specific cDNA clone CKFB10-E9.
 KW Cotton fibre; promoter; differential screening; leaf; ovule; root;
 KW flower; PCR; polymerase chain reaction; homology; transgenic plant; ds.
 OS Gossypium hirsutum.
 PN US521078-A.
 PD 28-MAY-1996.
 PF 04-OCT-1988; 253243.
 PR 04-OCT-1988; US-253243.
 PR 21-NOV-1990; US-617239.
 PR 18-MAY-1992; US-885970.
 PR 19-OCT-1994; US-298687.
 PA (CETU) AGRACETUS INC.
 PI John M;
 DR WPI: 96-267794/27.
 PT Isolation of fibre-specific cotton promoter sequences - using
 PT selected DNA probes to screen genomic DNA fragments, for production
 PT of cotton fibres with improved characteristics
 PS Examples; Column 45-46; 46pp; English.
 CC Cotton fibre cell-specific promoter sequences were isolated by
 CC differential screening of a cotton plant cDNA library. Of 4788 clones
 CC from a 10 day cell library screened with leaf cDNAs, 800 clones not
 CC present in the leaf were isolated. These were screened with cDNAs from
 CC ovule, root and flower mRNAs and resulted in 79 clones isolated. PCR
 CC analysis was then used to remove cross-hybridising clones. This resulted
 CC in the isolation of 18 cDNA clones specifically expressed in cotton fibre
 CC cells (T30242-4 and T30253-67). These cDNAs were then used to screen for
 CC homologous genomic sequences (T30245-53 and T30268) in order to obtain
 CC the corresp. promoter sequences.
 CC This cDNA clone contains an insert of 1283 bp which is highly expressed
 CC in fibres and weakly in petal tissues.
 CC The promoters isolated from the fibre cell-specific clones can be used to

CC generate transgenic cotton plants and lines producing fibres having
 CC altered quantity and quality.
 SQ Sequence 1283 BP; 509 A; 233 C; 251 G; 290 T;

alignment_scores:
 Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
 US-08-653-294-6 x T30265/rev ..

Align seg 1/1 to reverse of: T30265 from: 1 to: 1283

1 TyrArgLeuLeuIleArg 6
 |||||
 118 TACCGATTATTAGTGAGA 101

seq_name: N_Geneseq_36:T62624

seq_documentation_block:
 ID T62624 standard; cDNA to mRNA; 1283 BP.
 AC T62624;
 DT 14-MAY-1997 (first entry)
 DE Cotton fibre specific cDNA clone CKFB15-E9.
 KW cotton; fibre-specific; strength; transgenic plant; anthesis;
 KW developmentally regulated; E6; H6; antisense; sense; ss.
 OS Gossypium hirsutum strain Coker 312.
 PN US5597718-A.
 PD 28-JAN-1997.
 PF 04-OCT-1988; 253243.
 PR 04-OCT-1988; US-253243.
 PR 21-NOV-1990; US-617239.
 PR 18-OCT-1993; US-138814.
 PR 20-SEP-1995; US-530797.
 PA (CETU) AGRACETUS.
 PI Brill WJ, John ME, Umbeck PF;
 DR WPI: 97-108326/10.
 PT Prodn. of transgenic cotton plants - by transformation with the H6
 PT coding sequence or E6 anti-sense sequence, produces fibre of altered
 PT strength
 PS Example 4; Column 53-54; 33pp; English.
 CC T62609-24 are cotton fibre-specific cDNA clones which can be used to
 CC identify genomic clones. This clone, CKFB15-E9, is expressed in fibre
 CC cells, but is also expressed at low levels in petal. (CK -
 CC Coker; FB - Fibre; 10, 15 or 23 = age in days of fibre cells; Al and the
 CC last character and number stand for clone identity). The fibre-specific
 CC genes were identified by differential cDNA library screenings. Coding
 CC sequences from these isolated genes are used in sense or antisense
 CC orientation to alter the fibre characteristics, e.g. strength, of
 CC transgenic fibre-producing plants.
 SQ Sequence 1283 BP; 509 A; 233 C; 251 G; 290 T;

alignment_scores:
 Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
 US-08-653-294-6 x T62624/rev ..

Align seg 1/1 to reverse of: T62624 from: 1 to: 1283

1 TyrArgLeuLeuIleArg 6
 |||||
 118 TACCGATTATTAGTGAGA 101

seq_name: N_Geneseq_36:T70055

seq_documentation_block:
 ID T70055 standard; cDNA; 1283 BP.

AC T70055;
 DT 20-AUG-1997 (first entry)
 DE Cotton fibre specific cDNA clone E9.
 KW cotton; E9; fibre; promoter; transgenic plant; truncated;
 KW heterologous gene expression; ds.
 OS Gossypium hirsutum strain Coker 312.
 PN US5620882-A.
 PD 15-APR-1997.
 PF 04-OCT-1988; 253243.
 PR 04-OCT-1988; US-253243.
 PR 21-NOV-1990; US-617239.
 PR 18-MAY-1992; US-885970.
 PR 19-OCT-1994; US-298829.
 PA (CETU) AGRACETUS INC.
 PI John M;
 DR WPI; 97-235185/21.
 DT DNA constructs contg. truncated promoter sequence - for
 PT fibre-specific gene expression in cotton plants
 PS Example 3; Column 45-48; 48pp; English.
 CC T70040-57 are cotton fibre-specific cDNA clones which can be used to
 CC obtain genomic clones containing fibre-specific promoters. Claimed DNA
 CC constructs comprise a truncated promoter sequence (from one of T70031-38)
 CC that promotes preferential gene expression in plant fibre cells, a
 CC protein coding sequence not naturally associated with the promoter
 CC sequence and a 3' termination sequence. The DNA constructs are useful for
 CC expressing foreign genes in fibre-producing plants, esp. to produce
 CC transgenic cotton plants with varied cotton fibre characteristics and
 CC quality. The present sequence comprises E9 cDNA isolated from clone
 CC CKB15-E9 (CK = Coker; FB15 = 15 day old bolls).
 SQ Sequence 1283 BP; 509 A; 233 C; 251 G; 290 T;

alignment_scores:
 Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
 US-08-653-294-6 x T70055/rev ..

Align seg 1/1 to reverse of: T70055 from: 1 to: 1283

1 TyArgLeuLeuIleArg 6
 |||||
 118 TACCGATTATTAGTGA 101

seq_name: N_Geneseq_36:V44448

seq_documentation_block:
 ID V44448 standard; DNA; 2367 BP.

AC V44448;
 DT 09-NOV-1998 (first entry)
 DE Mycobacterium tuberculosis antigen RDIF7 DNA.
 KW Tuberculosis; infection; diagnosis; antigen; RDIF7; ss.
 OS Mycobacterium tuberculosis strain Erdman.
 FH Key Location/Qualifiers
 FT CDS 713..1843
 FT /*tag= a
 PN WO9816645-A2.
 PD 23-APR-1998.
 PF 07-OCT-1997; U18214.
 PR 13-MAR-1997; US-818111.
 PR 11-OCT-1996; US-729622.
 PA (CORI-) CORIXA CORP.
 PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ,
 PI Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
 DR WPI; 98-251292/22.
 DR P-PSDB: W64377.

PT New isolated Mycobacterium tuberculosis polypeptides and DNA - used
 PT to develop products for the detection of M. tuberculosis infection
 PT and diagnosis of tuberculosis
 PS Claim 11a; Page 207-209; 250pp; English.
 CC This DNA sequence codes for an antigenic portion (see W64377) of

CC Mycobacterium tuberculosis antigen RDIF7. It was isolated from a
 CC M. tuberculosis strain Erdman genomic DNA expression library using
 CC rabbit anti-sera raised against M. tuberculosis fractionated
 CC proteins. Phage plaques expressing immunoreactive antigens were
 CC purified. The invention relates to compositions and methods for
 CC diagnosing tuberculosis. It provides polypeptides (see W64291-
 CC W64379) comprising antigenic or immunogenic portions of M.
 CC tuberculosis antigens, as well as DNA sequences encoding such
 CC polypeptides, recombinant expression vectors and transformed or
 CC transfected host cells. Also claimed are methods and diagnostic
 CC kits for detecting M. tuberculosis infection in a patient using
 CC these polypeptides, antibodies or oligonucleotide probes and
 CC primers.
 SQ Sequence 2367 BP; 341 A; 1016 C; 583 G; 327 T;

alignment_scores:
 Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
 US-08-653-294-6 x V44448/rev ..

Align seg 1/1 to reverse of: V44448 from: 1 to: 2367

1 TyArgLeuLeuIleArg 6
 |||||
 1842 TATCGACTATTGTCGC 1825

seq_name: N_Geneseq_36:V64557

seq_documentation_block:
 ID V64557 standard; DNA; 2367 BP.

AC V64557;
 DT 27-JAN-1999 (first entry)
 DE M. tuberculosis immunogenic polypeptide RDIF7 DNA.
 KW Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
 KW vaccine; pharmaceutical; infection; diagnosis; ss.
 OS Mycobacterium tuberculosis.
 PN WO9816646-A2.
 PD 23-APR-1998.
 PF 07-OCT-1997; U18293.
 PR 13-MAR-1997; US-818112.
 PR 11-OCT-1996; US-730510.
 PA (CORI-) CORIXA CORP.
 PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ,
 PI Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
 DR WPI; 98-261042/23.
 DR P-PSDB: W81744.

PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used
 PT to develop products for the detection of M. tuberculosis infection
 PT and for diagnosis, treatment and prevention of tuberculosis
 PS Claim 4; Page 194-195; 230pp; English.
 CC This sequence encodes an immunogenic portion of a soluble Mycobacterium
 CC tuberculosis (MT) antigen which can be used in a method for inducing
 CC protective immunity against tuberculosis (TB). This sequence can be
 CC formulated into vaccines and/or pharmaceutical compositions for
 CC immunising against M. tuberculosis infection or may be used for the
 CC diagnosis of tuberculosis.
 SQ Sequence 2367 BP; 341 A; 1016 C; 583 G; 327 T;

alignment_scores:
 Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
 US-08-653-294-6 x V64557/rev ..

Align seg 1/1 to reverse of: V64557 from: 1 to: 2367

1 TyrArgLeuLeuIleArg 6
|||||
1842 TATCGACTATTGGTGGC 1825

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-6 to: EST:* out_format : pfs

Date: Feb 8, 2000 4:02 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet p2n.model -DEV=xlp
-O/cnrl_1/USPTO_spool/US08653294/runat_04022000_160700_15770/app_query.fasta.1
-DB-EST -QFMT=fastbp -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPT=0.000 -QGAPOP=4.500
-FGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-DELEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-LIST=7.000 -START=1 -MATRIX=blossum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294
-NCPU=6 -ICPU=3 -NO_XUPXY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-6

Query length: 6

Database: EST:*

Database sequences: 4538634

Database length: 1887831982

Search time (sec): 8553.360000

score_list:

| Sequence | Strd Orig | Zscore | Escore | Len | Documentation |
|-------------------|-----------|--------|--------|-----|----------------------------------|
| gb_est2:797779 | 29.00 | 132.71 | 125.96 | 70 | T97779 ye58h04.s1 Soares fetal 1 |
| gb_est15:AA470910 | 29.00 | 124.83 | 346.06 | 197 | AA470910 ne62a07.s1 NCI_CGAP.AL |
| gb_est17:AA621099 | 29.00 | 124.50 | 356.35 | 203 | AA621099 af34c01.s1 Soares tota |
| gb_est31:AI679465 | 29.00 | 124.38 | 366.63 | 209 | AI679465 tu74c04.x1 NCI_CGAP.Ga |
| gb_est19:AA751293 | 29.00 | 124.06 | 382.05 | 218 | AA751293 ISSU0105 Rice Immature |
| gb_est16:AA564601 | 29.00 | 123.95 | 387.18 | 221 | AA564601 nj02e06.s1 NCI_CGAP.Pd |
| gb_est14:AA274866 | 29.00 | 123.39 | 416.25 | 238 | AA274866 AV274866 RIKEN full-le |
| gb_est30:AV026198 | 29.00 | 122.98 | 438.44 | 251 | AV026198 AV026198 Mus musculus |
| gb_est42:AV285786 | 29.00 | 122.83 | 446.97 | 256 | AV285786 AV285786 RIKEN full-le |
| gb_est37:AV078270 | 29.00 | 122.51 | 465.72 | 267 | AV078270 AV078270 Rice green sh |
| gb_est39:AV143801 | 29.00 | 122.32 | 477.64 | 274 | AV143801 AV143801 RIKEN full-le |
| gb_est34:AV143803 | 29.00 | 122.04 | 494.66 | 284 | AV143803 AV143803 Mus musculus |
| gb_est20:AA872097 | 29.00 | 121.70 | 516.76 | 297 | AA872097 o112e04.s1 NCI_CGAP.GC |
| gb_gss1:CHS00983 | 29.00 | 121.63 | 521.86 | 300 | AL935841 Arabidopsis thaliana g |
| gb_est19:AA738876 | 29.00 | 120.72 | 586.33 | 338 | AA738876 vve2b01.r1 Soares 2NDM |
| gb_est15:H93368 | 29.00 | 120.67 | 589.72 | 340 | H93368 yw56h10.s1 Soares placen |
| gb_est30:AI642616 | 29.00 | 120.61 | 594.80 | 343 | AI642616 vve2b01.x1 Soares 2NDM |
| gb_gss8:AO61283 | 29.00 | 120.39 | 611.73 | 353 | AO61283 CIT-HSP-2349K21.TR CIT |
| gb_est17:C65958 | 29.00 | 120.24 | 623.58 | 360 | C65958 C65958 Yuji Kohara unpub |
| gb_est36:AV203011 | 29.00 | 120.24 | 623.58 | 360 | AV203011 AV203011 Yuji Kohara u |
| gb_est3:R69729 | 29.00 | 120.13 | 632.04 | 365 | R69729 y445e07.r1 Soares placen |
| gb_est20:AA872098 | 29.00 | 120.07 | 637.11 | 368 | AA872098 o112e05.s1 NCI_CGAP.GC |
| gb_est17:C69283 | 29.00 | 119.93 | 648.95 | 375 | C69283 C69283 Yuji Kohara unpub |
| gb_gss3:B82311 | 29.00 | 119.89 | 652.33 | 377 | B82311 RPI11-14717 TP RPI1-11 |
| gb_est14:AA448825 | 29.00 | 119.85 | 655.71 | 379 | AA448825 x10h06.s1 Soares tota |
| gb_est44:AW173025 | 29.00 | 119.85 | 655.71 | 379 | AW173025 x10h06.s1 Soares tota |
| gb_est23:AI076270 | 29.00 | 119.81 | 659.09 | 381 | AI076270 oy87c07.x1 Soares feta |
| gb_gss9:AO145722 | 29.00 | 119.63 | 674.29 | 390 | AO145722 HS_2216.A1 F09 MF CIT |
| gb_est38:AW058842 | 29.00 | 119.51 | 684.42 | 396 | AW058842 fe48h11.x1 Zebrafish W |
| gb_est15:AA58948 | 29.00 | 119.42 | 692.86 | 401 | AA58948 aa25h11.s1 NCI_CGAP.GC |
| gb_est24:AO029860 | 29.00 | 119.40 | 694.55 | 401 | AO029860 AO029860 Rice cDNA fr |
| gb_gss13:AI536554 | 29.00 | 119.38 | 696.23 | 403 | AI536554 t80e08.x1 NCI_CGAP.KI |
| gb_est13:AO468537 | 29.00 | 119.28 | 704.67 | 408 | AO468537 HS_2217.B2 C05_17A RPO |
| gb_gss10:AO221577 | 29.00 | 119.01 | 729.96 | 423 | AO221577 HS_2004.B2 C10 T7 CIT |
| gb_gss8:AO221027 | 29.00 | 118.83 | 746.81 | 433 | AO221027 HS_4756.A1 G09 SP6E CIT |
| gb_gss8:AO059883 | 29.00 | 118.83 | 746.81 | 433 | AO059883 CIT-HSP-2348L22.TR CIT |
| gb_gss8:AO061642 | 29.00 | 118.74 | 755.23 | 438 | AO061642 CIT-HSP-2348O11.TR CIT |
| gb_est38:AW025732 | 29.00 | 118.71 | 758.60 | 440 | AW025732 wu05g02.x1 NCI_CGAP.GC |
| gb_gss4:AO671287 | 29.00 | 118.66 | 763.65 | 443 | AO671287 HS_5452.A2 D04 SP6E RH |
| gb_gss12:AO355230 | 29.00 | 118.54 | 775.44 | 450 | AO355230 CITBI-EL-2534H15.TF CIT |
| gb_gss15:AO644952 | 29.00 | 118.54 | 775.44 | 450 | AO644952 RPI193-EcoRI-2C17.TP H |

gb_gss10:AO180569 + 29.00 118.50 778.80 452 ! AO180569 HS_3219.A2 H04_MR C
gb_gss5:AO790119 + 29.00 118.47 782.17 454 ! AO790119 HS_3247.A1_E05_MR C
gb_est39:AW097527 + 29.00 118.39 790.58 459 ! AW097527 rs45h03.y1 Sommer P

seq_name: gb_est2:T97779

seq_documentation_block:

LOCUS T97779 70 bp mRNA EST 29-MAR-1995
DEFINITION ye58h04.s1 Soares fetal liver spleen INFILS Homo sapiens cDNA clone
IMAGE:121975.3 similar to gb:M32011 NEUTROPHIL CYTOSOL FACTOR 2
(HUMAN); mRNA sequence.

ACCESSION T97779

VERSION T97779.1 GI:747124

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 70)

AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Huitman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Ruitkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevisakis,E., Waterston,R., Williamson,A., Wohldmann,P. and
Wilson,R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 2224

High quality sequence starts: 1 High quality sequence stops: 1

Source: IMAGE Consortium, LINL This clone is available royalty-free
through LINL; contact the IMAGE Consortium (info@image.lnl.gov)

for further information. Trace considered overall poor quality

Insert Length: 2224 Std error: 0.00

Seq primer: -21m3

High quality sequence stop: 1.

Location/Qualifiers

source

1. 70
/organism="Homo sapiens"
/db_xref="GDB:474520"
/db_xref="taxon:9606"
/clone="IMAGE:121975"
/clone_lib="Soares fetal liver spleen INFILS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH108 (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
15' AACGTGGAAGATTATTAAGACTTTTTTTTTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Patima Bonaldo."

BASE COUNT 29 a 10 c 10 g 21 t

ORIGIN

alignment_scores:

Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-6 x T97779/rev ..

Align seg 1/1 to reverse of: T97779 from: 1 to: 70

```

1 TtArgLeuLeuIleArg 6
|||||
50 TATAGGTACTCATAGG 33

seq_name: gb_est15:AA470910

seq_documentation_block:
LOCUS AA470910 197 bp mRNA EST 14-AUG-1997
DEFINITION ne62a07.s1 NCI_CGAP_Alvl Homo sapiens cDNA clone IMAGE:908820, mRNA
sequence.
ACCESSION AA470910
VERSION AA470910.1 GI:2198219
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 197)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1394316.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 364 Std Error: 0.00
Seq primer: -4lm13 fwd. ET from Amersham.
Location/Qualifiers
1. 197
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:908820"
/clone_lib="NCI_CGAP_Alvl"
/tissue_type="alveolar rhabdomyosarcoma"
/lab_host="DH10B"
/notes="Vector: PAMP10; mRNA made from alveolar
rhabdomyosarcoma, cDNA made by oligo-dT priming.
Non-directionally cloned. Size-selected on agarose gel,
average insert size 600 bp. Reference: Krizman et al.
(1996) Cancer Research 56:5380-5383."
BASE COUNT 62 a 40 C 50 g 45 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x AA470910 ..
Align seg 1/1 to: AA470910 from: 1 to: 197

1 TtArgLeuLeuIleArg 6
|||||
105 TACAGACTACTCATAGG 122

seq_name: gb_est17:AA621099

seq_documentation_block:
LOCUS AA621099 203 bp mRNA EST 02-MAR-1998
DEFINITION af34c01.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
sequence.
ACCESSION AA621099
VERSION AA621099.1 GI:2525038
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 203)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1407372.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 841 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. 203
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1033536"
/clone_lib="Soares_total_fetus_Nb2HF8_9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCTTAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Ronaldo."
BASE COUNT 70 a 32 c 32 g 69 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x AA621099 ..
Align seg 1/1 to: AA621099 from: 1 to: 203

1 TtArgLeuLeuIleArg 6
|||||
27 TACAGACTACTCATCGT 44

seq_name: gb_est31:AI679465

seq_documentation_block:
LOCUS AI679465 209 bp mRNA EST 26-MAY-1999
DEFINITION tu74c04.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2256774 3'
similar to gb:MI4219 BONE PROTEOGLYCAN II PRECURSOR (HUMAN);, mRNA
sequence.
ACCESSION AI679465

```


VERSION AI679465.1 GI:4889647
 KEYWORDS EST
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 209)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Mar 10, 1998 this sequence version replaced gi:2948836.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Trace considered overall poor quality
 Seq primer: -40UP from Gibco
 High quality sequence stop: 1.
 Location/Qualifiers
 FEATURES
 source
 1..209
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2256774"
 /clone_lib="NCI-CGAP_Gas4"
 /tissue_type="poorly differentiated adenocarcinoma with
 signed ring cell features"
 /lab_hosts="DH10B"
 /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.69 kb. Life Technologies catalog #:
 11549-011"
 BASE COUNT 70 a 47 c 38 g 54 t
 ORIGIN
 alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-08-653-294-6 x AI679465/rev ..
 Align seg 1/1 to reverse of: AI679465 from: 1 to: 209
 1 TyrArgLeuLeuLeuArg 6
 |||||
 201 TACAGGCTCTATTAGG 184
 seq_name: gb_est19:AA751293
 seq_documentation_block:
 LOCUS AA751293 218 bp mRNA EST 20-JAN-1998
 DEFINITION ISSU0105 Rice Immature Seed Lambda ZAPII CDNA Library Oryza sativa
 CDNA clone ISSU0105, mRNA sequence.
 ACCESSION AA751293
 VERSION AA751293.1 GI:2797999
 KEYWORDS EST.
 SOURCE Oryza sativa.
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
 Poaceae; Oryza.
 REFERENCE 1 (bases 1 to 218)

AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
 Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
 Lee,M.C. and Eun,M.Y.
 TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
 JOURNAL Unpublished (1998)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1404701.
 Contact: Eun M.Y.
 Department of Cyto genetics
 National Inst. of Agri. Sci. and Tech, RDA
 Suwon, Kyunggido, Korea
 Tel: 82 331 290 0301
 Fax: 82 331 290 0307
 Email: myeun@sun20.asti.re.kr
 Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
 University, Yongin, Korea. 449-728 bhnahm@bserver.myongji.ac.kr
 Seq primer: M13 Reverse Primer.
 Location/Qualifiers
 source
 1..218
 /organism="Oryza sativa"
 /cultivar="Milyang23"
 /db_xref="taxon:4530"
 /clone="ISSU0105"
 /clone_lib="Rice Immature Seed Lambda ZAPII CDNA Library"
 /tissue_type="Immature Seed"
 /dev_stage="5 days after pollination"
 /lab_hosts="E. coli SOLR"
 /note="Vector: Bluescript SK(+); Site_1: EcoRI; Site_2:
 XhoI; Directional CDNA library inserted into lambda ZAPII
 vector at 5' end with EcoRI and 3' end with Xho I site."
 BASE COUNT 65 a 45 c 45 g 61 t 2 others
 ORIGIN
 alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-08-653-294-6 x AA751293/rev ..
 Align seg 1/1 to reverse of: AA751293 from: 1 to: 218
 1 TyrArgLeuLeuLeuArg 6
 |||||
 173 TATAGACTACTCATAGA 156
 seq_name: gb_est16:AA564601
 seq_documentation_block:
 LOCUS AA564601 221 bp mRNA EST 04-SEP-1997
 DEFINITION nj02806.s1 NCI-CGAP_Pr21 Homo sapiens CDNA clone IMAGE:985186 3'
 similar to gb:M27492 INTERLEUKIN-1 RECEPTOR, TYPE I PRECURSOR
 (HUMAN);, mRNA sequence.
 ACCESSION AA564601
 VERSION AA564601.1 GI:2336240
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 221)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Dec 18, 1996 this sequence version replaced gi:1734341.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution Information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 797 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES

```

1. .221
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:985186"
/clone_lib="NCI_CGAP_Pr21"
/sex="male"
/tissue_type="normal prostate"
/lab_host="DH10b"
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; 1st strand cDNA was prepared
from normal prostate bulk tissue, and was then primed with
a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is not normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

```

BASE COUNT
ORIGIN

```
alignment_scores:
  Quality: 29.00      Length: 6
  Ratio: 4.833      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:

Align seq 1/1 to: AA564601 from: 1 to: 221

```

1 TyrArgLeuLeuIleArg 6
  |||||
24 TACAGGCTATTGATCCGT 41

```

seq_name: qb_est41:AV274866

| seq_documentation_block: | | |
|--------------------------|---|-------------|
| LOCUS | 238 bp | EST |
| AV274866 | RIKEN full-length enriched, | 05-NOV-1999 |
| DEFINITION | musculus cDNA clone 4932422P17 3', mRNA sequence. | |

| | |
|-----------|-------------|
| ACCESSION | AV274866 |
| VERSION | AV274866.1 |
| | GI:52622903 |

KEYWORDS
SOURCE
EST.
HOUSE MOUSE

ORGANISM

REFERENCE

AUTHORS

Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Horii, F., Ishii, Y., Ishikawa, T., Itoh, M., Iwama, M., Kadota, K., Kasaga, I., Kat, C., Kawai, J., Kikuchi, N., Kojima, T., Koya, S., Kusabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tominga, N., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y. (KONNO, H. et al.)

TITLE

RIKEN MOUSE ESTs (KONNO, H. et al.)

TITLE

JOURNAL

COMMENT On Apr 30, 1999 this sequence version replaced gi:4727909.

Contact: Yoshihide Hayashizaki

Genome Exploration Research Group, Life Science Tsukuba Center,
Genome Science Laboratory
The Institute of Physical and Chemical Research (RIKEN), Genomic
Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: +81-298-36-9013
Fax: +81-298-36-9098
Email: genome-res@rtc.riken.go.jp,
sarak@rtc.riken.go.jp,
usuaki.n@genome.rtc.riken.go.jp,
Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,
Matsuura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and
Hayashizaki,Y.

Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 95 (7), 3455-3460 (1998)

Itoch, M., Kitesunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303, 19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

source
1. .238

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="4932422P1"
/clone_lib="RIKEN full-length enriched, adult male testis
(DH10B)"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trihalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGGAGAGATTTTCAGGTAAATTAATATATCCCCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."

```

**BASE . CO
ORIGIN**

alignment scores:

| | | | |
|-------------|---------|-------------------|---------|
| Quality: | 29.00 | Length: | 6 |
| Ratio: | 4.833 | Gaps: | 0 |
| Similarity: | 100.000 | Percent Identity: | 100.000 |

alignment block:

US-08-653-294-6 X AV274866/rev

Align seq 1/1 to reverse of: AV274866 from: 1 to: 238

1 TyrArgLeuLeuIleArg 6

— — — — —

82 TATAGGCTAATCCGT 65

me: gb_est30:AV026198

cumentation_block:

EST.
house mouse.
Mus musculus
Eukaryota: Metazoa: Chordata: Vertebrata: Mammalia:
Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mmus.
1 (bases 1 to 256)
Konno,H., Azawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,
Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I.,
Kai,C., Kawaj,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,
Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K.,
Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y.,
Suzuki,H., Suzuki,H., Takanashi,F., Tatenno,M., Tomiyaga,N.,
Tsunoda,Y., Watanishi,A., Watanabe,S., Yamamura,T., Yasunishi,A.,
Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Riken Mouse ESTs (Konno,H., et al., Unpublished (1999))
On Dec 20, 1995 this sequence version replaced gi:1133578.

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: +81-298-36-9013
Fax: +81-298-36-9098
E-mail: genome-res@rtc.riken.go.jp,
URL: <http://genome.rtc.riken.go.jp/>
Sasaki, N., Izawa, M., Wataniki, M., Ozawa, K., Tanaka, T., Yoneda, Y.,
Matsushima, S., Carninci, P., Muramatsu, M., Okazaki, Y. and
Hayashizaki, Y.
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 95 (7), 3455-3460 (1998)
Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,
Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. *Genome Res.* 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303,
19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp/>) for
further details.

Okazaki, Y. and Hayashizaki, Y. Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303, 19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

Location/Qualifiers

1. .256

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="5031414A18"

/clone_lib="RIKEN full-length enriched, 11 days pregnant adult female ovary and uterus"

/sex="female"

/tissue_type="ovary and uterus"
 /dev_stage="11 days pregnant, adult"
 /lab_host="DH10B"
 /note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGATCGATCCGCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot=10.0 and subtraction to Rot=100.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATCGATCGATTAATAATCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plucuescript KS(+), after bulk excision from Lambda Fx1. Cloning sites, 5' end: SalI 3' end:

```

BamHI."
BASE COUNT      69 a  48 c  35 g  104 t
ORIGIN
alignment_scores:
  Quality:      29.00      Length:      6
  Ratio:        4.833      Gaps:        0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x AV258786/rev ..
Align seg 1/1 to reverse of: AV258786 from: 1 to: 256

1 TyrArgLeuLeulleArg 6
|||||
25 TATCGTTTGCTCATTCGA 8

seq_name: gb_est37:AU078270

seq_documentation_block:
LOCUS      AU078270      267 bp      mRNA      EST      08-SEP-1999
DEFINITION AU078270 Rice green shoot Oryza sativa cDNA clone S12251_122, mRNA
sequence.
ACCESSION  AU078270
VERSION    AU078270.1 GI:5851098
KEYWORDS   EST.
SOURCE     Oryza sativa.
  ORGANISM  Oryza sativa
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
    Poaceae; Oryza.
REFERENCE  1 (bases 1 to 267)
AUTHORS   Yamamoto,K. and Sasaki,T.
TITLE     Rice cDNA from green shoot
JOURNAL   Unpublished (1996)
COMMENT   On May 18, 1998 this sequence version replaced gi:3137393.
          Contact: Takuji Sasaki
          National Institute of Agrobiological Resources
          Rice Genome Research Program
          2-1-2 Kannondai,Tsukuba
          Ibaraki,
          Japan 305
          Tel: 0298-38-7441
          Fax: 0298-38-7468
          Email: tsasaki@agr.affrc.go.jp
          PROJECT = "RGP".

FEATURES             Location/Qualifiers
     source           1..267
                     /organism="Oryza sativa"
                     /strain="Nipponbare"
                     /db_xref="taxon:4530"
                     /clone="S12251_122"
                     /clone_lib="Rice green shoot"
                     /note="Green shoot (8 days old)"

BASE COUNT      80 a  39 c  51 g  87 t
ORIGIN
alignment_scores:
  Quality:      29.00      Length:      6
  Ratio:        4.833      Gaps:        0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x AU078270 ..
Align seg 1/1 to: AU078270 from: 1 to: 267

1 TyrArgLeuLeulleArg 6
|||||
20 TACCGTTTGTTGATTAGG 37

```

seq_name: gb_est39:AV218201

seq_documentation_block: 274 bp mRNA EST 30-OCT-1999
 LOCUS AV218201 RIKEN full-length enriched, adult male hippocampus Mus
 DEFINITION musculus cDNA clone 2900087E03 3', mRNA sequence.

ACCESSION AV218201
 VERSION AV218201.1 GI:6159042

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 274)

AUTHORS

Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
 Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,
 Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I.,
 Kai,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,
 Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
 Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K.,
 Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Suganara,Y.,
 Suzuki,H., Suzuki,H., Takahashi,F., Tateno,M., Tomimaga,N.,
 Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A.,
 Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Konno,H., et al.)

TITLE

JOURNAL

COMMENT

On Dec 20, 1995 this sequence version replaced gi:1134368.

Contact: Yoshihide Hayashizaki

Genome Exploration Research Group, Life Science Tsukuba Center,

The Institute of Physical and Chemical Research (RIKEN), Genomic

Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: +81-298-36-9013

Fax: +81-298-36-9098

Email: genome-res@rtc.riken.go.jp,

URL:http://genome.rtc.riken.go.jp/

Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,

Matsuura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and

Hayashizaki,Y.

Transcriptional sequencing: A method for DNA sequencing using RNA

polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)

Itch,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,

Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,

Okazaki,Y. and Hayashizaki,Y.

Automated filtration-based high-throughput plasmid preparation

system. Genome Res. 9 (5), 463-470 (1999)

Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,

19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for

further details.

FEATURES

source

Location/Qualifiers

1..274

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="2900087E03"

/clone_lib="RIKEN full-length enriched, adult male

hippocampus"

/sex="male"

/tissue_type="hippocampus"

/dev_stage="adult"

/lab_host="SOLR"

/note="Site 1: XhoI; Site 2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in

RIKEN, Division of Experimental Animal Research in Riken

contributed to prepare mouse tissues. 1st strand cDNA was

primed with a primer [5'

GAGACAGAGAGGATCCAGAGCTCTTTTITTTTTTIVN 3'), cDNA was

prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5',
GAGAGAGATTCGAGTTAAATTAATCCGCCCCCCCCC 3']"

BASE COUNT 57 a 87 c 46 g 84 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-6 x AV218201/rev ..
Align seg 1/1 to reverse of: AV218201 from: 1 to: 274

1 TyArgLeuLeuIleArg 6
|||||

62 TATAGGCTTCGATTAGG 45

seq_name: gb_est34:AV143803

seq_documentation_block:
LOCUS AV143803 284 bp mRNA EST 02-JUL-1999
DEFINITION AV143803 Mus musculus C57BL/6J 10-11 day embryo Mus musculus cDNA
clone 2810432E15, mRNA sequence.

ACCESSION AV143803

VERSION AV143803.1 GI:5347798

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

AUTHORS

Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,
Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T.,
Hara, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M.,
Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Niitsuma, H., Oda, H.,
Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y.,
Sugahara, Y., Suzuki, H., Suzuki, H., Tateno, M., Tonari, Y.,
Tomihaga, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T.,
Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

TITLE RIKEN Mouse ESTs

JOURNAL

COMMENT

On May 18, 1998 this sequence version replaced gi:3136837.

Contact: Chile Owa

Genome Science Laboratory

RIKEN

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-9145

Fax: 81-298-36-9098

Email: genome-resetc.riken.go.jp

Thermotabilization and thermoactivation of thermostable enzymes by
trehalose and its application for the synthesis of full length cDNA
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))

Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))

Please visit our web site (<http://genome.rtc.riken.go.jp>) for
further details.

location/Qualifiers

1. 284

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="2810432E15"

/clone_lib="Mus musculus C57BL/6J 10-11 day embryo"

/sex="mixed"

/dev_stage="10-11 day embryo"

BASE COUNT 62 a 80 c 43 g 99 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-6 x AV143803 ..

Align seg 1/1 to: AV143803 from: 1 to: 284

1 TyArgLeuLeuIleArg 6

|||||

82 TACCGCTTCCTATTCGA 99

seq_name: gb_est20:AA872097

seq_documentation_block:

LOCUS AA872097 297 bp mRNA EST 17-MAR-1998
DEFINITION O112804.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1476318 3'
similar to gb:M32011 NEUTROPHIL CYTOSOL FACTOR 2 (HUMAN);, mRNA
sequence.

ACCESSION AA872097

VERSION AA872097.1 GI:2968275

KEYWORDS EST.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Sep 12, 1996 this sequence version replaced gi:1404743.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Seq primer: -40ml3 fwd. RT from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1. 297

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1476318"

/clone_lib="NCI_CGAP_GC4"

/tissue_type="pooled germ cell tumors"

/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; 1st strand cDNA was prepared from 3 pooled

germ cell tumors, and was then primed with a Not I -

oligo(dT) primer. Double-stranded cDNA was ligated to Eco

RI adaptors (Pharmacia), digested with Not I and cloned

into the Not I and Eco RI sites of the modified pT73

vector. Library is normalized. Library was constructed by

Bento Soares and M. Fatima Bonaldo.

BASE COUNT 93 a 54 c 67 g 83 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x AA872097/rev ..
Align seg 1/1 to reverse of: AA872097 from: 1 to: 297

1 TyrArgLeuLeuIleArg 6
|||||
50 TATAGGTACTCATAGG 33

seq_name: gb_gssl:CNS00Y83

seq_documentation_block:
LOCUS CNS00Y83 300 bp DNA GSS 29-JUN-1999
DEFINITION Arabidopsis thaliana genome survey sequence sp6 end of BAC T15N20
of TAMU library from strain Columbia of Arabidopsis thaliana,
genomic survey sequence.
ACCESSION AL095841
VERSION AL095841.1 GI:5303996
KEYWORDS GSS.
SOURCE
ORGANISM Arabidopsis thaliana
thale cress.
REFERENCE 1 (bases 1 to 300)
AUTHORS Salanoubat,M., Choisne,N., Artiguenave,F., Brottier,P., Wincker,P.,
Samson,D., Saurin,W., Weissenbach,J. and Quetier,F.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 300)
AUTHORS Direct Submission
TITLE Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage :
JP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
JOURNAL - Web : www.genoscope.cns.fr
FEATURES
source
1..300
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_lib="TAMU"
/clone="T15N20"
/note="end : SP6"
BASE COUNT 47 a 92 c 87 g 74 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x CNS00Y83/rev ..
Align seg 1/1 to reverse of: CNS00Y83 from: 1 to: 300

1 TyrArgLeuLeuIleArg 6
|||||
266 TACCGGCTCTGATCCGC 249

seq_name: gb_est19:AA738876

seq_documentation_block:
LOCUS AA738876 338 bp mRNA EST 14-JAN-1998
DEFINITION vv62b01.r1 Soares 2NbMT Mus musculus cDNA clone IMAGE:1226953 5',
mRNA sequence.
ACCESSION AA738876
VERSION AA738876.1 GI:2776128
KEYWORDS EST.

SOURCE
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS 1 (bases 1 to 338)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilton,R. and
Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:652545
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 337.
FEATURES
source
1..338
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1226953"
/clone_lib="Soares 2NbMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCGTGTGTGTGTGTGTGTGTGTGT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT 44 a 102 c 126 g 66 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-6 x AA738876/rev ..
Align seg 1/1 to reverse of: AA738876 from: 1 to: 338

1 TyrArgLeuLeuIleArg 6
|||||
269 TACCGGCTCTCATCGG 252

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:29:36 ; Search time 122.56 Seconds
(without alignments)
1.160 Million cell updates/sec

Title: US-08-653-294-7

Perfect score: 29

Sequence: 1 YRLAIR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 29 | 100.0 | 6 | W47263 | Immunomodulatory p |
| 2 | 29 | 100.0 | 6 | W33781 | Peptide #2 used in |
| 3 | 29 | 100.0 | 10 | W47266 | Immunomodulatory p |
| 4 | 29 | 100.0 | 10 | W47270 | Immunomodulatory p |
| 5 | 29 | 100.0 | 12 | R95429 | HLA-B2702 84-79-84 |
| 6 | 29 | 100.0 | 12 | W33798 | Peptide B2702 84-7 |
| 7 | 29 | 100.0 | 12 | W33799 | Immunomodulating d |
| 8 | 29 | 100.0 | 20 | R92909 | HLA-B2702 CTL modu |
| 9 | 29 | 100.0 | 20 | R92911 | HLA-B2702 CTL modu |
| 10 | 29 | 100.0 | 20 | R92907 | HLA-B2702 CTL modu |
| 11 | 29 | 100.0 | 20 | R95428 | HLA-B2702 84-75-84 |
| 12 | 29 | 100.0 | 20 | R95430 | HLA-B2702 84-75T/7 |
| 13 | 29 | 100.0 | 20 | W33778 | Immunomodulating d |
| 14 | 29 | 100.0 | 20 | W33779 | Immunomodulating d |
| 15 | 29 | 100.0 | 20 | W33792 | Peptide B2702 84-7 |
| 16 | 28 | 96.6 | 347 | W72010 | HSV-2 strain SB5 C |
| 17 | 28 | 96.6 | 450 | W72145 | HSV-2 strain SB5 C |
| 18 | 28 | 96.6 | 548 | W72165 | HSV-2 strain SB5 C |
| 19 | 26 | 89.7 | 255 | W71313 | Helicobacter polyp |
| 20 | 25 | 86.2 | 43 | R58902 | Drosophila-12 cadh |
| 21 | 25 | 86.2 | 43 | R87142 | Protocadherin clon |
| 22 | 25 | 86.2 | 783 | W27113 | Rat spleen pro-hor |
| 23 | 24 | 82.8 | 6 | W47264 | Immunomodulatory p |
| 24 | 24 | 82.8 | 6 | W33783 | Peptide #4 used in |
| 25 | 24 | 82.8 | 10 | W47268 | Immunomodulatory p |
| 26 | 24 | 82.8 | 10 | W47272 | Immunomodulatory p |
| 27 | 24 | 82.8 | 20 | R92910 | HLA-B2702 CTL modu |
| 28 | 24 | 82.8 | 20 | R92908 | HLA-B2702 CTL modu |
| 29 | 24 | 82.8 | 20 | W33791 | Peptide B2702 84-7 |
| 30 | 24 | 82.8 | 20 | W33793 | Peptide B2702 84-7 |
| 31 | 24 | 82.8 | 162 | W98864 | H. pylori GHPO 167 |
| 32 | 24 | 82.8 | 243 | P70483 | Sequence encoded b |
| 33 | 24 | 82.8 | 248 | P60437 | Dog 32 kd alveolar |
| 34 | 24 | 82.8 | 256 | R04210 | Canine 32K alveola |

Human thioresoxin
KM31-7 precursor.
Human KM-102-deriv
Bacillus thuringie
Enzyme M-11. DNA e
Enzyme Q36. DNA en
Human thioresoxin
Human thioresoxin
Manduca sexta Bac1
Manduca sexta BT t
Mouse DEC-205. L1g

35 24 82.8 504 1 W83401
36 24 82.8 549 1 R92050
37 24 82.8 551 1 W83404
38 24 82.8 724 1 R93081
39 24 82.8 772 1 R79949
40 24 82.8 775 1 R79950
41 24 82.8 939 1 W83402
42 24 82.8 939 1 W83403
43 24 82.8 1528 1 R95333
44 24 82.8 1528 1 W90182
45 24 82.8 1723 1 W00645

ALIGNMENTS

RESULT 1
W47263
ID W47263 standard; peptide; 6 AA.
AC W47263;
DT 22-MAY-1998 (first entry)
DE Immunomodulatory peptide.
KW Immunomodulator; Class I HLA-B alpha-1 domain; Inhibition;
KW transplant rejection; treatment; autoimmune disease.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 1..6 "at least one of the amino acids is the D-isomer"
FT W09744052-A1.
PN 27-NOV-1997.
PF 23-APR-1997; U06705.
PR 22-MAY-1996; US-651650.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM;
DR WPI; 98-018220/02.
PT Novel immunomodulatory peptide-type compound - useful for inhibiting transplant rejection
PS Claim 10; Page 36; 41pp; English.
CC The present sequence is an immunomodulatory peptide, which comprises a Class I HLA-B alpha-1 domain sequence. It can be used in a pharmaceutical composition together with a subtherapeutic dose of an immunosuppressant, to extend the period of acceptance of a CC transplant from a major histocompatibility complex (MHC) unmatched CC donor, i.e. to inhibit transplant rejection. It can also be used in CC the treatment of autoimmune diseases.
CC Peptides using the D-form amino acids are more effective
CC Immunomodulators than their diastereomers or enantiomers.
SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
Db 1 YRLAIR 6

RESULT 2
W33781
ID W33781 standard; peptide; 6 AA.
AC W33781;
DT 19-JUN-1998 (first entry)
DE Peptide #2 used in immunomodulating dimer peptide.
KW Immunomodulating dimer; Immunosuppressant drug; CTL activation;
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain; rejection.
OS Synthetic.
OS Homo sapiens.
PN W09744351-A1.

27-NOV-1997. U08689.
 22-MAY-1997; US-653294.
 24-MAY-1996; US-653294.
 (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI: 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Claim 15; Page 35; 41pp; English.
 CC This sequence represents a specifically claimed peptide which forms part
 CC of the immunomodulating dimer peptides of the invention. A peptide-type
 CC compound or variant is claimed which has immunomodulating activity,
 CC including the N-terminal acylated and/or C-terminal amidated or
 CC esterified forms of up to 60 amino acids, where the peptide-type compound
 CC comprises the formula: A-B, where A, B = (R aa76-77L) (aa79-84) or
 CC (aa84-79) (Laa77-76R); aa76 = E or V; aa77 = D, S or N; aa79 = R or G;
 CC aa80 = I or N; aa81, aa84 = a hydrophobic or small amino acid; aa82 = R
 CC or L; aa83 = G or R; and aa represents amino acid. The sequence in the
 CC brackets may optionally be absent or truncated at any peptide type bond
 CC within the brackets. The compounds comprise amino acid sequences related
 CC to a Class I HLA-B alpha1 domain (positions 79-84). They can be used to
 CC inhibit cytotoxic T-lymphocytes (CTL) from undesirably attacking cells in
 CC a host or in vitro. They can also be used in combination with antigenic
 CC peptides or proteins of interest to activate CTLs. They can also inhibit
 CC the proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.
 SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
 |||||
 Db 1 YRLAIR 6

RESULT 3

W47266
 ID W47266 standard; peptide; 10 AA.
 AC W47266;
 DT 22-MAY-1998 (first entry)
 DE Immunomodulatory peptide.
 KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
 KW transplant rejection; treatment; autoimmune disease.
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc_difference 1..10
 FT /note= "at least one of the amino acids is the
 FT D-isomer

Query Match 100.0%; Score 29; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
 |||||
 Db 1 YRLAIR 6

RESULT 5

R95429
 ID R95429 standard; peptide; 12 AA.
 AC R95429;
 DT 12-NOV-1996 (first entry)
 DE HLA-B*2702 84-79-84 palindromic.
 KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KW cytolysis; antigen presenting cell.
 OS Synthetic.
 PN W09513288-Al.
 PD 18-MAY-1995.
 PF 10-NOV-1994; U12985.
 PR 10-NOV-1993; US-150493.

CC immunomodulators than their diastereomers or enantiomers.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.25;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
 |||||
 Db 1 YRLAIR 6

RESULT 4

W47270
 ID W47270 standard; peptide; 10 AA.
 AC W47270;
 DT 22-MAY-1998 (first entry)
 DE Immunomodulatory peptide.
 KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
 KW transplant rejection; treatment; autoimmune disease.
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc_difference 1..10
 FT /note= "at least one of the amino acids is the
 FT D-isomer

PN W09744052-Al.
 PD 27-NOV-1997.
 PF 23-APR-1997; U06705.
 PR 22-MAY-1996; US-651650.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI: 98-018220/02.
 PT Novel immunomodulatory peptide-type compound - useful for inhibiting
 PT transplant rejection
 PS Claim 10; Page 36; 41pp; English.
 CC The present sequence is an immunomodulatory peptide, which
 CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
 CC in a pharmaceutical composition together with a subtherapeutic dose
 CC of an immunosuppressant, to extend the period of acceptance of a
 CC transplant from a major histocompatibility complex (MHC) unmatched
 CC donor, i.e. to inhibit transplant rejection. It can also be used in
 CC the treatment of autoimmune diseases.
 CC Peptides using the D-form amino acids are more effective
 CC immunomodulators than their diastereomers or enantiomers.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 29; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.25;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
 |||||
 Db 1 YRLAIR 6

PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI; 95-194027/25.
 PT Compens. comprising lymphoid surface membrane proteins - which may
 PT inhibit cytolytic activity and differentiation of CTLs.
 PS Example; Page 12; 29pp; English.
 CC R95413, and R95415-R95431 represent palindromes and fragments of
 CC human-leucocyte-associated antigens. This sequence represents the
 CC HLA-B2702 84-79-84 palindrome. These sequences can be used to isolate
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
 CC protein associated with T-cell activation in mammalian T-cells, and is
 CC also immunologically cross reactive with the heat shock protein Hsc70.
 CC p74 is found in a limited number of cell types, but is particularly
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable
 CC cell with an amphoteric detergent, and then passed through an affinity
 CC column containing a covalently bound HLA-B2702 palindromic peptide.
 CC Compositions comprising the extracellular fragment of p74 combined with
 CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits
 CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate
 CC compounds can be screened for their effect on the cytolytic activity of
 CC T-cells, by combining them with the extracellular portion of p74 and
 CC determining the amount of binding between the candidate compound and p74.
 CC Modulation of CTL activity can be inhibited in a cellular composition
 CC containing T-cells and antigen presenting cells (APCs), by adding to the
 CC mix the extracellular portion of p74, in an amount sufficient to compete
 CC with p74 for the binding of the p74 ligand.
 SQ Sequence 12 AA;

Query Match 100.0%; Score 29; DB 1; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.3;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRLAIR 6
 DB 1 YRLAIR 6

RESULT 7
 W33798
 ID W33798 standard; peptide; 12 AA.
 AC W33798;
 DT 19-JUN-1998 (first entry)
 DE Peptide B2702.84-79/79-84 tested for immunomodulating activity.
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN WO9744351-A1.
 PD 27-NOV-1997.
 PF 22-MAY-1997; U08689.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI; 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Claim 17; Page 35; 41pp; English.
 CC This sequence represents a specifically claimed immunomodulating
 CC dimer peptide of the invention. A peptide-type compound or variant is
 CC claimed which has immunomodulating activity, including the N-terminal
 CC acylated and/or C-terminal amidated or esterified forms of up to 60
 CC amino acids, where the peptide-type compound comprises the formula: A-B,
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa
 CC represents amino acid. The sequence in the brackets may optionally be
 CC absent or truncated at any peptide type bond within the brackets. The
 CC compounds comprise amino acid sequences related to a Class I HLA-B
 CC alpha1 domain (positions 79-84). They can be used to inhibit cytotoxic
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in
 CC vitro. They can also be used in combination with antigenic peptides or
 CC proteins of interest to activate CTLs. They can also inhibit the
 CC proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.
 SQ Sequence 12 AA;

Query Match 100.0%; Score 29; DB 1; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.3;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRLAIR 6
 DB 1 YRLAIR 6

RESULT 8
 R92909

```

ID R92909 standard; peptide; 20 AA.
AC R92909; 1996 (first entry)
DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84(T)).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW class I MHC; HLA-B2702.
OS Synthetic.
PN WO9526979-A1.
PD 12-OCT-1995.
PF 05-APR-1995; U04349.
PR 05-APR-1994; US-222851.
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM, Parham P;
DR WPI; 95-358582/46.
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
PS Example 15; Page 36; 80pp; English.
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of
CC class I major histocompatibility complex (MHC) antigens. This sequence
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class
CC I MHC HLA-B2702. These sequences can be used to extend the period of
CC acceptance by a recipient of a transplant from an MHC unmatched donor.
CC The peptides are administered to a patient in conjunction with a
CC subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.
SQ Sequence 20 AA;

Query Match 100.0%; Score 29; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
DB 1 YRLAIR 6

RESULT 9
ID R92911 standard; peptide; 20 AA.
AC R92911; 1996 (first entry)
DE HLA-B2702 CTL modulating peptide (B2702.84-75/84-75).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW class I MHC; HLA-B2702.
OS Synthetic.
PN WO9526979-A1.
PD 12-OCT-1995.
PF 05-APR-1995; U04349.
PR 05-APR-1994; US-222851.
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM, Parham P;
DR WPI; 95-358582/46.
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
PS Example 15; Page 36; 80pp; English.
CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of
CC class I major histocompatibility complex (MHC) antigens. This sequence
CC is a dimer of residues 75-84 of the alpha-1 domain of the class I MHC
CC HLA-B2702. These sequences can be used to extend the period of
CC acceptance by a recipient of a transplant from an MHC unmatched donor.
CC The peptides are administered to a patient in conjunction with a
CC subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.
SQ Sequence 20 AA;

Query Match 100.0%; Score 29; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
DB 1 YRLAIR 6

RESULT 9
ID R92911 standard; peptide; 20 AA.
AC R92911; 1996 (first entry)
DE HLA-B2702 CTL modulating peptide (B2702.84-75/84-75).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW class I MHC; HLA-B2702.
OS Synthetic.
PN WO9526979-A1.
PD 12-OCT-1995.
PF 05-APR-1995; U04349.
PR 05-APR-1994; US-222851.
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM, Parham P;
DR WPI; 95-358582/46.
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
PS Example 15; Page 36; 80pp; English.
CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of
CC class I major histocompatibility complex (MHC) antigens. This sequence
CC is a dimer of residues 75-84 of the alpha-1 domain of the class I MHC
CC HLA-B2702. These sequences can be used to extend the period of
CC acceptance by a recipient of a transplant from an MHC unmatched donor.
CC The peptides are administered to a patient in conjunction with a
CC subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.
SQ Sequence 20 AA;

Query Match 100.0%; Score 29; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
DB 1 YRLAIR 6

RESULT 10
ID R92907 standard; peptide; 20 AA.
AC R92907; 1996 (first entry)
DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW class I MHC; HLA-B2702.
OS Synthetic.
PN WO9526979-A1.
PD 12-OCT-1995.
PF 05-APR-1995; U04349.
PR 05-APR-1994; US-222851.
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM, Parham P;
DR WPI; 95-358582/46.
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
PS Example 15; Page 36; 80pp; English.
CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of
CC class I major histocompatibility complex (MHC) antigens. This sequence
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class
CC I MHC HLA-B2702. These sequences can be used to extend the period of
CC acceptance by a recipient of a transplant from an MHC unmatched donor.
CC The peptides are administered to a patient in conjunction with a
CC subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.
SQ Sequence 20 AA;

Query Match 100.0%; Score 29; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
DB 1 YRLAIR 6

RESULT 11
ID R95428 standard; peptide; 20 AA.
AC R95428; 1996 (first entry)
DE HLA-B2702 84-75-84 palindromic.
KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
KW cytolysis; antigen presenting cell.
OS Synthetic.
PN WO9513288-A1.
PD 18-MAY-1995.
PR 10-NOV-1994; U12985.
PA (STRD ) UNIV LELAND STANFORD JUNIOR.

```

PI Clayberger C, Krensky AM;
 DR WPI; 95-194027/25.
 PT Compsns. comprising lymphoid surface membrane proteins - which may
 PT inhibit cytolytic activity and differentiation of CTLs.
 PS Example; Page 12; 29pp; English.
 CC R95413, and R95415-R95431 represent palindromes and fragments of
 CC human-leucocyte-associated antigens. This sequence represents the
 CC HLA-B2702 84-75-84 palindrome. These sequences can be used to isolate
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
 CC protein associated with T-cell activation in mammalian T-cells, and is
 CC also immunologically cross reactive with the heat shock protein Hsc70.
 CC p74 is found in a limited number of cell types, but is particularly
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable
 CC cell with an amphoteric detergent, and then passed through an affinity
 CC column containing a covalently bound HLA-B2702 palindromic peptide.
 CC Compositions comprising the extracellular fragment of p74 combined with
 CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits
 CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate
 CC compounds can be screened for their effect on the cytolytic activity of
 CC T-cells, by combining them with the extracellular portion of p74 and
 CC determining the amount of binding between the candidate compound and p74.
 CC Modulation of CTL activity can be inhibited in a cellular composition
 CC containing T-cells and antigen presenting cells (APCs). By adding to the
 CC mix the extracellular portion of p74, in an amount sufficient to compete
 CC with p74 for the binding of the p74 ligand.
 SQ Sequence 20 AA;

Query Match 100.0%; Score 29; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.52;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRLAIR 6
 DB 1 YRLAIR 6

RESULT 12

ID R95430 standard; peptide; 20 AA.
 AC R95430;
 DT 12-NOV-1996 (first entry)
 DE HLA-B2702 84-75T/75-84T palindrome.
 KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KW cytotoxic; antigen presenting cell.
 OS Synthetic.
 PN W09513288-Al.
 PD 18-MAY-1995.
 PR 10-NOV-1994; U12985.
 PR 10-NOV-1993; US-150493.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI; 95-194027/25.
 PT Compsns. comprising lymphoid surface membrane proteins - which may
 PT inhibit cytolytic activity and differentiation of CTLs.
 PS Example; Page 12; 29pp; English.
 CC R95413, and R95415-R95431 represent palindromes and fragments of
 CC human-leucocyte-associated antigens. This sequence represents the
 CC HLA-B2702 84-75T/75-84T palindrome. These sequences can be used to
 CC isolate the protein p74 from a T-cell lysate. p74 is a T-cell surface
 CC membrane protein associated with T-cell activation in mammalian T-cells,
 CC and is also immunologically cross reactive with the heat shock protein
 CC Hsc70. p74 is found in a limited number of cell types, but is
 CC particularly expressed on B and T cells. p74 can be isolated by lysis of
 CC a suitable cell with an amphoteric detergent, and then passed through an
 CC affinity column containing a covalently bound HLA-B2702 palindromic
 CC peptide. Compositions comprising the extracellular fragment of p74
 CC combined with HLA-B2702.60-84 (see R95416), induces calcium influx, and
 CC inhibits cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity.
 CC Candidate compounds can be screened for their effect on the cytolytic
 CC activity of T-cells, by combining them with the extracellular portion of

CC p74 and determining the amount of binding between the candidate compound
 CC and p74. Modulation of CTL activity can be inhibited in a cellular
 CC composition containing T-cells and antigen presenting cells (APCs), by
 CC adding to the mix the extracellular portion of p74, in an amount
 CC sufficient to compete with p74 for the binding of the p74 ligand.
 SQ Sequence 20 AA;

Query Match 100.0%; Score 29; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.52;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRLAIR 6
 DB 1 YRLAIR 6

RESULT 13

W33778
 ID W33778 standard; peptide; 20 AA.
 AC W33778;
 DT 19-JUN-1998 (first entry)
 DE Immunomodulating dimer peptide #1.
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN W09744351-Al.
 PD 27-NOV-1997.
 PR 22-MAY-1997; U08689.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI; 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Claim 16; Page 35; 41pp; English.
 CC This sequence represents a specifically claimed immunomodulating
 CC dimer peptide of the invention. A peptide-type compound or variant is
 CC claimed which has immunomodulating activity, including the N-terminal
 CC acylated and/or C-terminal amidated or esterified forms of up to 60
 CC amino acids, where the peptide-type compound comprises the formula: A-B,
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-78R); aa76 = E or
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa
 CC represents amino acid. The sequence in the brackets may optionally be
 CC absent or truncated at any peptide type bond within the brackets. The
 CC compounds comprise amino acid sequences related to a Class I HLA-B
 CC alpha1 domain (positions 79-84). They can be used to inhibit cytotoxic
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in
 CC vitro. They can also be used in combination with antigenic peptides or
 CC proteins of interest to activate CTLs. They can also inhibit the
 CC proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.
 SQ Sequence 20 AA;

Query Match 100.0%; Score 29; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.52;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRLAIR 6
 DB 1 YRLAIR 6

RESULT 14

W33779
 ID W33779 standard; peptide; 20 AA.

AC W33779;
 DT 19-JUN-1998 (first entry)
 DE Immunomodulating dimer peptide #2.
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN WO9744351-A1.
 PD 27-NOV-1997. U08689.
 PF 22-MAY-1997; US-653294.
 PR (STRD) UNIV LELAND STANFORD JUNIOR.
 PA Beulow R, Clayberger C, Krensky AM;
 PI WPI; 98-086530/08.
 DR New Immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PT Claim 16; Page 35; 41pp; English.
 CC This sequence represents a specifically claimed immunomodulating
 CC dimer peptide of the invention. A peptide-type compound or variant is
 CC claimed which has immunomodulating activity, including the N-terminal
 CC acylated and/or C-terminal amidated or esterified forms of up to 60
 CC amino acids, where the peptide-type compound comprises the formula: A-B,
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa
 CC represents amino acid. The sequence in the brackets may optionally be
 CC absent or truncated at any peptide type bond within the brackets. The
 CC compounds comprise amino acid sequences related to a Class I HLA-B
 CC alpha domain (positions 79-84). They can be used to inhibit cytotoxic
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in
 CC vitro. They can also be used in combination with antigenic peptides or
 CC proteins of interest to activate CTLs. They can also inhibit the
 CC proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.
 SQ Sequence 20 AA;

Query Match 100.0%; Score 29; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.52;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
 | | | | |
 Db 1 YRLAIR 6

RESULT 15
 W33792
 ID W33792 standard; peptide; 20 AA.
 AC W33792;
 DT 19-JUN-1998 (first entry)
 DE Peptide B2702.84-75/75-84 tested for immunomodulating activity.
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN WO9744351-A1.
 PD 27-NOV-1997.
 PF 22-MAY-1997; U08689.
 PR (STRD) UNIV LELAND STANFORD JUNIOR.
 PA Beulow R, Clayberger C, Krensky AM;
 PI WPI; 98-086530/08.
 DR New Immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PT Claim 1; Page 19; 41pp; English.
 PS Peptides W33784-98 and W33778-9 were assayed for their immunomodulating

CC activity. A peptide-type compound or variant is claimed which has
 CC immunomodulating activity, including the N-terminal acylated and/or
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where
 CC the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes, be
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 SQ Sequence 20 AA;

Query Match 100.0%; Score 29; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.52;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
 | | | | |
 Db 1 YRLAIR 6

Search completed: February 8, 2000, 01:29:37
 Job time: 1749 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 11:54:18 ; Search time 117.7 seconds
(without alignments)
2.405 Million cell updates/sec

Title: US-08-653-294-7
Perfect score: 29
Sequence: 1 YRLAIR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : PIR_62:*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 29 | 100.0 | 1124 | 2 D65032 | hypothetical prote |
| 2 | 28 | 96.6 | 388 | 2 S15593 | hypothetical prote |
| 3 | 28 | 96.6 | 579 | 2 A7367 | oligopeptide ABC t |
| 4 | 27 | 93.1 | 185 | 2 S7446 | hypothetical prote |
| 5 | 27 | 93.1 | 309 | 2 G70882 | probable oxidoredu |
| 6 | 27 | 93.1 | 404 | 2 F71324 | probable gcpE prot |
| 7 | 26 | 89.7 | 96 | 2 B69327 | conserved hypothet |
| 8 | 26 | 89.7 | 105 | 2 T10473 | molt-inhibiting ho |
| 9 | 26 | 89.7 | 255 | 2 H71963 | fumarate reductase |
| 10 | 26 | 89.7 | 235 | 2 A64544 | fumarate reductase |
| 11 | 26 | 89.7 | 291 | 2 F72660 | hypothetical prote |
| 12 | 26 | 89.7 | 545 | 2 JN0448 | t-complex polypept |
| 13 | 26 | 89.7 | 883 | 2 S31175 | hypothetical prote |
| 14 | 25 | 86.2 | 148 | 2 B69960 | 3-dehydroquinatase |
| 15 | 25 | 86.2 | 190 | 2 B42957 | iron-sulfur protei |
| 16 | 25 | 86.2 | 195 | 2 F64955 | filZ protein - Esc |
| 17 | 25 | 86.2 | 223 | 2 D49804 | capsid protein p25 |
| 18 | 25 | 86.2 | 258 | 2 S10164 | fumarate reductase |
| 19 | 25 | 86.2 | 368 | 2 F72281 | hypothetical prote |
| 20 | 25 | 86.2 | 380 | 2 H70590 | hypothetical prote |
| 21 | 25 | 86.2 | 381 | 2 T06293 | 3-hydroxyisobutyry |
| 22 | 25 | 86.2 | 463 | 2 B72500 | probable seryl-trn |
| 23 | 25 | 86.2 | 545 | 2 T00485 | probable phosphori |
| 24 | 25 | 86.2 | 629 | 2 S60385 | probable membrane |
| 25 | 25 | 86.2 | 744 | 2 T10035 | hypothetical prote |
| 26 | 25 | 86.2 | 747 | 2 D70802 | hypothetical prote |
| 27 | 25 | 86.2 | 783 | 2 JC6136 | keatin-like protein |
| 28 | 25 | 86.2 | 785 | 2 S64706 | subtilisin-like pr |
| 29 | 25 | 86.2 | 907 | 2 T16911 | hypothetical prote |
| 30 | 25 | 86.2 | 1603 | 1 BVASA1 | arom protein - Eme |

ALIGNMENTS

RESULT 1

D65032
hypothetical protein b2549 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 14-Nov-1997
C:Accession: D65032
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: D65032
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1124 <BLAT>
A:Cross-references: GB:AE000341; GB:U00096; NID:g1788899; PID:g1788900; UWGP:b2549
A:Experimental source: strain K-12, substrain MG1655

Query Match 100.0%; Score 29; DB 2; Length 1124;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
|||||
DB 412 YRLAIR 417

RESULT 2

S15593
hypothetical protein (insertion sequence ISH27-3) - Halobacterium halobium
C:Species: Halobacterium halobium
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 30-Jun-1998
C:Accession: S15593
R:Pfeifer, F.; Blaseio, U.
Nucleic Acids Res. 18, 6921-6925, 1990
A:Title: Transposition burst of the ISH27 insertion element family in Halobacterium h
A:Reference number: S15591; MUID:91088266
A:Accession: S15593
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-388 <PPE>
A:Cross-references: EMBL:X54434
A:Note: the authors translated the initiation codon GTG for residue 1 as Val
C:Genetics:
A:Mobile element: insertion sequence ISH27-3
A:Start codon: GTG

Query Match 96.6%; Score 28; DB 2; Length 388;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
 Db 154 YRLAVR 159
 RESULT 3
 A72367
 oligopeptide ABC transporter, periplasmic oligopeptide-binding protein - Thermotoga maritima
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
 C:Accession: A72367
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing
 A:Reference number: A72200; MUID:99287316
 A:Accession: A72367
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-579 <ARN>
 A:Cross-references: GB:AE001728; GB:AE000512; NID:g4981027; PID:g4981044; TIGR:TM0531
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM0531

Query Match 96.6%; Score 28; DB 2; Length 579;
 Best Local Similarity 83.3%; Pred. No. 23;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YRLAIR 6
 Db 168 YRLAVR 173

RESULT 4
 S74416
 hypothetical protein sil0687 - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 21-Aug-1998
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
 A:Reference number: S74322; MUID:97061201
 A:Accession: S74416
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-185 <KAN>
 A:Cross-references: EMBL:D64001; GB:AB001339; NID:gl001102; PID:dl010985; PID:gl001190
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 93.1%; Score 27; DB 2; Length 185;
 Best Local Similarity 83.3%; Pred. No. 13;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YRLAIR 6
 Db 35 YRLAIR 40

RESULT 5
 G70882
 probable oxidoreductase - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 29-Sep-1999
 C:Accession: G70882
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

: Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Nature 393, 537-544, 1998
 A:Authors: Krogh, A.; Mclean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Qua.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987
 A:Accession: G70882
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-309 <COL>
 A:Cross-references: GB:AL008967; GB:AL123456; NID:g3261491; PIDN:CAA15591.1; PID:el29
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: Rv2776c
 C:Superfamily: phthalate dioxygenase reductase; cytochrome-b5 reductase homology; fer
 F9-211/Domain: cytochrome-b5 reductase homology <CBR>
 F240-297/Domain: ferredoxin [2Fe-2S] homology <FER>

Query Match 93.1%; Score 27; DB 2; Length 309;
 Best Local Similarity 83.3%; Pred. No. 22;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YRLAIR 6
 Db 65 YRLAIR 70

RESULT 6
 F71324
 probable gcpe protein (gcpe) - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 17-Mar-1999
 C:Accession: F71324
 R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; M
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
 A:Reference number: A71250; MUID:98332170
 A:Accession: F71324
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-404 <COL>
 A:Cross-references: GB:AE001221; GB:AE000520; NID:g3322720; PID:g3322731
 A:Experimental source: strain Nichols
 C:Genetics:
 A:Gene: TP0446

Query Match 93.1%; Score 27; DB 2; Length 404;
 Best Local Similarity 83.3%; Pred. No. 29;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YRLAIR 6
 Db 121 YRLAIR 126

RESULT 7
 B69327
 conserved hypothetical protein AF0618 - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-Jun-1998
 C:Accession: B69327
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
 : Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes,
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
 A:Reference number: A69250; MUID:98049343

A:Accession: B69327
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-96 <KLE>
A:Cross-references: GB:AE001062; GB:AE000782; NID:g2689385; PID:g2650002; TIGR:AF0618

Query Match 89.7%; Score 26; DB 2; Length 96;
Best Local Similarity 83.3%; Pred. No. 13;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
|||||
DB 51 YRLAIR 56

RESULT 8

molit-inhibiting hormone precursor - Penaeus japonicus
C:Species: Penaeus japonicus
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T10473
R:Ohira, T.; Watanabe, T.; Nagasawa, H.; Aida, K.
Zool. Sci. 14, 785-789, 1997
A:Title: Molecular cloning of a molit-inhibiting hormone cDNA from the kuruma prawn Penaeus japonicus
A:Reference number: 217038
A:Accession: T10473
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-105 <OHI>
A:Cross-references: EMBL:AB004652; NID:d1107174; PID:d1021266
A:Keywords: hormone; neuropeptide
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-105/Product: molit-inhibiting hormone #status predicted <MAT>

Query Match 89.7%; Score 26; DB 2; Length 105;
Best Local Similarity 83.3%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
|||||
DB 2 YRLAIR 7

RESULT 9

fumarate reductase - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 11-Jun-1999
C:Accession: H71963
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori
A:Reference number: A71800; MUID:99120557
A:Accession: H71963
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-255 <ARN>
A:Cross-references: GB:AE001456; GB:AE001439; NID:g4154689; PIDN:AND05763.1; PID:g4154689
A:Experimental source: strain J99
C:Genetics:
A:Gene: frdC
C:Superfamily: fumarate reductase cytochrome b

Query Match 89.7%; Score 26; DB 2; Length 255;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
|||||
DB 188 YRLAIR 193

RESULT 10
A64544
fumarate reductase, cytochrome b subunit - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 11-Jun-1999
C:Accession: A64544
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R. Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKee, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, A.; Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467
A:Accession: A64544
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-255 <TOM>
A:Cross-references: GB:AE000539; GB:AE000511; NID:g2313275; PIDN:AAD07260.1; PID:g2313275
C:Superfamily: fumarate reductase cytochrome b

Query Match 89.7%; Score 26; DB 2; Length 255;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
|||||
DB 188 YRLAIR 193

RESULT 11

F72660
hypothetical protein APE0710 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: F72660
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Tawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J. DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero-thermus aquaticus
A:Reference number: A72450; MUID:99310339
A:Accession: F72660
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-291 <RAW>
A:Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAW79686.1; PID:d1043472; PID:g5104188
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0710

Query Match 89.7%; Score 26; DB 2; Length 291;
Best Local Similarity 83.3%; Pred. No. 37;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
|||||
DB 66 YRLAIR 71

RESULT 12

JN0448
t-complex polypeptide Tcpl-1 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Aug-1999
C:Accession: JN0448
R:Mori, M.; Murata, K.; Kubota, H.; Yamamoto, A.; Matsushiro, A.; Morita, T. Gene 122, 381-382, 1992

A:Title: Cloning of a cDNA encoding the Tcp-1 (t complex polypeptide 1) homologue of *Ar*
 A:Reference number: JN0448; MUID:93138412
 A:Accession: JN0448
 A:Molecule type: mRNA
 A:Residues: 1-545 <NR>
 A:CROSS-references: DDBJ:DI1351; NID:g217870; PIDN:BAA01955.1; PID:di002434; PID:g217871
 C:Superfamily: molecular chaperone t-complex-type

Query Match 89.7%; Score 26; DB 2; Length 545;
 Best Local Similarity 83.3%; Pred. No. 69;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
 |::|::|
 Db 124 YRLAIR 129

RESULT 13

S31175
 hypothetical protein 2 - midge (*Chironomus thummi*) transposon NLR1Cth
 C:Species: *Chironomus thummi*
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 29-Jan-1999
 C:Accession: S31175
 R:Blinov, A.G.; Sobanov, Y.V.; Bogachev, S.S.; Donchenko, A.P.; Filippova, M.A.
 Mol. Gen. Genet. 237, 412-420, 1993
 A:Title: The *Chironomus thummi* genome contains a non-LTR retrotransposon.
 A:Reference number: S31174; MUID:93247556
 A:Accession: S31175
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-883 <BLI>
 A:CROSS-references: GB:S59870; NID:g299960; PID:g299962

Query Match 89.7%; Score 26; DB 2; Length 883;
 Best Local Similarity 83.3%; Pred. No. 11e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
 |::|::|
 Db 803 YRLAIR 808

RESULT 14

B69960
 3-dehydroquinolate dehydratase homolog yqhs - *Bacillus subtilis*
 C:Species: *Bacillus subtilis*
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Jun-1999
 C:Accession: B69960
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 C.; Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 y, M.; Ogawa, K.; Ogiwara, C.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69580; MUID:98044033
 A:Accession: B69960
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-148 <KUN>
 A:CROSS-references: GB:299116; GB:AL009126; NID:g2634723; PIDN:CAB14378.1; PID:el185715;
 A:Experimental source: strain 168
 C:Genetics:

A:Gene: yqhs
 C:Superfamily: catabolic 3-dehydroquinolate dehydratase

Query Match 86.2%; Score 25; DB 2; Length 148;
 Best Local Similarity 66.7%; Pred. No. 35;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
 |::|::|
 Db 132 YRLAIR 137

RESULT 15

B42957
 iron-sulfur protein Coof - *Rhodospirillum rubrum*
 C:Species: *Rhodospirillum rubrum*
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 11-Jun-1999
 C:Accession: B42957
 R:Kerby, R.L.; Hong, S.S.; Ensign, S.A.; Coppoc, L.J.; Ludden, P.W.; Roberts, G.P.
 J. Bacteriol. 174, 5284-5294, 1992
 A:Title: Genetic and physiological characterization of the *Rhodospirillum rubrum* carb
 A:Reference number: A42957; MUID:92355502
 A:Accession: B42957
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-190 <KER>
 A:CROSS-references: GB:U65510; GB:U43789; GB:M90421; GB:S41762; GB:U20508; NID:g15154
 A:Experimental source: URL
 A>Note: sequence extracted from NCBI backbone (NCBI:110353, NCBIP:110359)
 C:Superfamily: nrfC protein; ferredoxin 2[4Fe-4S] homology
 C:Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein
 F:10-85/Domain: ferredoxin 2[4Fe-4S] homology <PER1>
 F:89-162/Domain: ferredoxin 2[4Fe-4S] homology <PER2>
 F:17,20,23,77/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted
 F:27,65,68,73/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted
 F:96,99,102,154/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted
 F:106,133,136,150/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 86.2%; Score 25; DB 2; Length 190;
 Best Local Similarity 66.7%; Pred. No. 44;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
 |::|::|
 Db 167 YRLAIR 172

Search completed: February 7, 2000, 11:54:20
 Job time: 24330 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 00:59:48 ; Search time 63.71 Seconds
(without alignments)
2.813 Million cell updates/sec

Title: US-08-653-294-7

Perfect score: 29

Sequence: 1 YRLAIR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SwissProt_38.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 29 | 100.0 | 1124 | 1 YPHG_ECOLI | P76585 escherichia |
| 2 | 27 | 93.1 | 230 | 1 CBBY_RHOSH | P95649 rhodobacter |
| 3 | 27 | 93.1 | 470 | 1 PABE_LACLA | P27629 lactococcus |
| 4 | 26 | 89.7 | 105 | 1 MIH_PENJP | P55847 pentacoccus |
| 5 | 26 | 89.7 | 255 | 1 FRDC_HELPY | O06912 helicobacte |
| 6 | 26 | 89.7 | 545 | 1 TCPA_ARATH | P28769 arabidopsis |
| 7 | 25 | 86.2 | 148 | 1 3DHQ_BACSU | P54517 bacillus su |
| 8 | 25 | 86.2 | 183 | 1 FLI2_ECOLI | P52627 escherichia |
| 9 | 25 | 86.2 | 190 | 1 COOF_RHORI | P31894 rhodospirill |
| 10 | 25 | 86.2 | 223 | 1 COAT_CTV36 | O00686 citrus tris |
| 11 | 25 | 86.2 | 256 | 1 FRDC_WOLSU | P17413 wolfinella s |
| 12 | 25 | 86.2 | 629 | 1 FRE7_YEAST | Q12333 saccharomyc |
| 13 | 25 | 86.2 | 1586 | 1 AROL_EHANI | P07547 e pentafunc |
| 14 | 24 | 82.8 | 113 | 1 YF65_SYNY3 | P74596 synchocyst |
| 15 | 24 | 82.8 | 123 | 1 YG2E_YEAST | P53247 saccharomyc |
| 16 | 24 | 82.8 | 156 | 1 RM25_YEAST | P23369 saccharomyc |
| 17 | 24 | 82.8 | 243 | 1 IPT2_AGR7 | P06524 agrobacteri |
| 18 | 24 | 82.8 | 243 | 1 IPT_AGRRA | P14011 agrobacteri |
| 19 | 24 | 82.8 | 248 | 1 PSPA_CANFA | P06908 canis famil |
| 20 | 24 | 82.8 | 248 | 1 PSPA_MOUSE | P35242 mus musculu |
| 21 | 24 | 82.8 | 249 | 1 PSPA_PIG | P49874 sus scrofa |
| 22 | 24 | 82.8 | 271 | 1 RL5_DUNSA | O22608 dunaliella |
| 23 | 24 | 82.8 | 294 | 1 RL5A_SCHPO | P52822 schizosacch |
| 24 | 24 | 82.8 | 294 | 1 RL5B_SCHPO | O74306 schizosacch |
| 25 | 24 | 82.8 | 297 | 1 ARCC_ECOLI | P73706 escherichia |
| 26 | 24 | 82.8 | 297 | 1 RL5_HELAN | O65353 helianthus |
| 27 | 24 | 82.8 | 297 | 1 RL5_YEAST | P26321 saccharomyc |
| 28 | 24 | 82.8 | 301 | 1 RL5_NEUCR | O59953 neurospora |
| 29 | 24 | 82.8 | 306 | 1 SYV_METJA | O57834 methanococc |
| 30 | 24 | 82.8 | 315 | 1 NMPP_ECOLI | P34210 escherichia |
| 31 | 24 | 82.8 | 349 | 1 NTRB_ECOLI | P06712 escherichia |
| 32 | 24 | 82.8 | 349 | 1 NTRB_KLEPN | P06218 klebsiella |
| 33 | 24 | 82.8 | 349 | 1 NTRB_SALT | P41788 salmonella |
| 34 | 24 | 82.8 | 366 | 1 DP3B_BUCAP | P29439 buchnera ap |

| | | | | | |
|----|----|------|-----|--------------|--------------------|
| 35 | 24 | 82.8 | 373 | 1 OPS1_DROME | P06002 drosophila |
| 36 | 24 | 82.8 | 374 | 1 OPS1_DROPS | P28678 drosophila |
| 37 | 24 | 82.8 | 376 | 1 DP3B_STRCO | P27903 streptomyc |
| 38 | 24 | 82.8 | 377 | 1 OPS2_HEMSA | Q25158 hemigrapsus |
| 39 | 24 | 82.8 | 397 | 1 DP3B_MYCSM | P52851 mycobacteri |
| 40 | 24 | 82.8 | 399 | 1 DP3B_MYCLE | P46387 mycobacteri |
| 41 | 24 | 82.8 | 402 | 1 DP3B_MYCTU | Q50790 mycobacteri |
| 42 | 24 | 82.8 | 408 | 1 YS92_MYCTU | Q10813 mycobacteri |
| 43 | 24 | 82.8 | 446 | 1 AP50_SCHPO | Q09718 schizosacch |
| 44 | 24 | 82.8 | 494 | 1 PACE_BPP1 | P27753 bacterioph |
| 45 | 24 | 82.8 | 497 | 1 TRXB_HUMAN | Q16881 homo sapien |

ALIGNMENTS

RESULT 1
YPHG_ECOLI STANDARD; PRT; 1124 AA.
AC P76585;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHETICAL 127.3 KD PROTEIN IN CSIE-GLYA INTERGENIC REGION.
GN YPHG.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE; 97426617.
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., ROSE C.K., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000341; AAC75602.1; -
DR ECOGENE; EGI3468; yphg.
KW Hypothetical protein.
SQ SEQUENCE 1124 AA; 127284 MW; 0A06B4C6 CRC32;

Query Match 100.0%; Score 29; DB 1; Length 1124;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
Db 412 YRLAIR 417

RESULT 2
CBBY_RHOSH STANDARD; PRT; 230 AA.
ID CBBY_RHOSH
AC P95649;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CBBY PROTEIN.

OS Rhodobacter sphaeroides (Rhodopseudomonas sphaeroides).
GN Rhodobacter sphaeroides; alpha subdivision; Rhodobacter group;
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;

OC Rhodobacter.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HR;
 RX MEDLINE: 97158658.
 RA GIBSON J.L., TABITA F.R.;
 RT "Analysis of the cbxYZ operon in Rhodobacter sphaeroides.";
 RL J. Bacteriol. 179:663-669(1997).
 CC -!- SIMILARITY: BELONGS TO THE CBXY/CBBZ/GPH/YIEH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U67781; AAC44828.1;
 DR PFAM: PF00702; Hydrolase; 1;
 SQ SEQUENCE 230 AA; 25118 MW; 54A55A59 CRC32;

Query Match 93.1%; Score 27; DB 1; Length 230;
 Best Local Similarity 83.3%; Pred. No. 8.9;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRLAIR 6
 DB 157 YRLAIR 162
 |||||

RESULT 3
 ID PABE_LACLA STANDARD; PRT; 470 AA.
 AC P27629;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE PARA-AMINO-BENZOATE SYNTHASE COMPONENT I (EC 4.1.3.-) (ADC SYNTHASE).
 GN PABE.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Lactococcus.
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=NCDO 496;
 RX MEDLINE: 94014976.
 RA ARHIN F.F., VINING L.C.;
 RT "Cloning, nucleotide sequence and expression in Streptomyces lividans
 RT and Escherichia coli of pabB from Lactococcus lactis subsp. lactis
 RT NCDO 496.";
 RL J. Gen. Microbiol. 139:1785-1793(1993).
 CC -!- FUNCTION: CATALYZES THE BIOSYNTHESIS OF 4-AMINO-4-DEOXYCHORISMATE
 CC (ADC) FROM CHORISMATE AND GLUTAMINE.
 CC -!- PATHWAY: FOLATE BIOSYNTHESIS PATHWAY. FIRST STEP IN THE
 CC BIOSYNTHESIS OF P-AMINO-BENZOATE (PABA).
 CC -!- SUBUNIT: CONSISTS OF TWO NONIDENTICAL CHAINS: COMPONENT I
 CC CATALYZES THE FORMATION OF ADC BY BINDING CHORISMATE AND AMMONIA;
 CC COMPONENT II PROVIDES THE GLUTAMINE AMIDOTRANSFERASE ACTIVITY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M64860; AAA17025.1;
 DR PFAM: PF00425; chorismate_bind; 1.
 KW Lyase; Folate biosynthesis.
 SQ SEQUENCE 470 AA; 50970 MW; 41AC304B CRC32;

Query Match 93.1%; Score 27; DB 1; Length 470;
 Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRLAIR 6
 DB 411 YRLAIR 416
 |||||

RESULT 4
 ID MIH_PENJP STANDARD; PRT; 105 AA.
 AC P5847; O02379;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE MOLT-INHIBITING HORMONE PRECURSOR (MIH) (PEI-SGP-IV).
 OS Penaeus japonicus (Kuruma prawn).
 CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 CC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeidae;
 CC Penaeus.
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=EYESTALK;
 RX MEDLINE: 98112189.
 RA OHIRA T., WATANABE T., NAGASAWA H., AIDA K.;
 RT "Molecular cloning of a molt-inhibiting hormone cDNA from the kuruma
 RT prawn Penaeus japonicus.";
 RL Zool. Sci. 14:785-789(1997).
 [2]
 RN SEQUENCE OF 29-105.
 RP TISSUE=SINUS GLAND;
 RX MEDLINE: 9625356.
 RA YANG W.-J., AIDA K., TERAUCHI A., SONOBE H., NAGASAWA H.;
 RT "Amino acid sequence of a peptide with molt-inhibiting activity from
 RT the kuruma prawn Penaeus japonicus.";
 RL Peptides 17:197-202(1996).
 CC -!- FUNCTION: INHIBITS Y-ORGANS WHERE MOLTING HORMONE (ECDYSTEROID) IS
 CC SECRETED. A MOLTING CYCLE IS INITIATED WHEN MIH SECRETION
 CC DIMINISHES OR STOPS. HAS LITTLE OR NO HYPERGLYCEMIC ACTIVITY.
 CC -!- TISSUE SPECIFICITY: PRODUCED BY THE MEDULLA TERMINALIS X-ORGAN IN
 CC THE EYESTALKS AND TRANSPORTED TO THE SINUS GLAND WHERE IT IS
 CC STORED AND RELEASED.
 CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD CHH/MIH/GIH/VIH FAMILY OF
 CC HORMONES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB004652; BAA20432.1;
 DR PROSITE: PS01250; CHH_MIH_GIH; 1.
 DR PFAM: PF01147; Crust_neurohorm; 1.
 KW Signal; Neuropeptide; Hormone.
 [1]
 RN SIGNAL 28
 FT PEPTIDE 29 105 MOLT-INHIBITING HORMONE.
 FT DISULFID 35 72 BY SIMILARITY.
 FT DISULFID 52 68 BY SIMILARITY.
 FT DISULFID 55 81 BY SIMILARITY.
 FT CONFLICT 44 45 IY -> YN (IN REF. 2).
 SQ SEQUENCE 105 AA; 12150 MW; CA4F62D8 CRC32;

Query Match 89.7%; Score 26; DB 1; Length 105;
 Best Local Similarity 83.3%; Pred. No. 7;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 YRLAIR 6
    ||||:
Db 2 YRLAMR 7

RESULT 5
FRDC_HELPY
ID FRDC_HELPY STANDARD; PRT; 255 AA.
AC O06912;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FUMARATE REDUCTASE CYTOCHROME B SUBUNIT.
GN FRDC OR HP0193.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-U0802 / NCTC 11639;
RA GE 2., JIANG Q., KALISIAK M.S., TAYLOR D.E.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-26695 / ATCC 700392;
RX MEDLINE: 97394467.
RA TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,
RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,
RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,
RA LORTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,
RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,
RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,
RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WAYTHEY L., WALLIN E.,
RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,
RA VENTER J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -!- FUNCTION: DI-HEME CYTOCHROME OF THE FUMARATE REDUCTASE COMPLEX (BY
CC SIMILARITY).
CC -!- SUBUNIT: FUMARATE DEHYDROGENASE FORMS PART OF AN ENZYME COMPLEX
CC CONTAINING THREE SUBUNITS: A FLAVOPROTEIN, AN IRON-SULFUR, AND
CC CYTOCHROME B-556 (BY SIMILARITY).
CC -!- SIMILARITY: TO CYTOCHROME B-558 FROM B.SUBILIS SUCCINATE
CC DEHYDROGENASE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U78101; AAC46063.1; -.
CC DR EMBL; AE000539; AAD07260.1; -.
CC DR TIGR; HP0193; -.
CC KW Tricarboxylic acid cycle; Transmembrane; Electron transport; Heme.
CC TRANSMEM 33 53
CC FT TRANSMEM 78 98
CC FT TRANSMEM 126 146
CC FT TRANSMEM 168 188
CC FT TRANSMEM 208 228
CC FT BINDING 44 44
CC FT BINDING 93 93
CC FT BINDING 143 143
CC FT BINDING 182 182
CC FT CONFLICT 86 86
CC FT CONFLICT 90 90
CC FT CONFLICT 200 200
CC FT CONFLICT 236 236
CC FT CONFLICT 255 AA; 28886 MW; AEC224C CRC32;
CC SEQUENCE 255 AA; 28886 MW; AEC224C CRC32;

OY 1 YRLAIR 6
    ||||:
Db 2 YRLAMR 7

Query Match 89.7%; Score 26; DB 1; Length 255;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YRLAIR 6
    ||||:
Db 188 YRLAIK 193

RESULT 6
TCPA_ARATH
ID TPCA_ARATH STANDARD; PRT; 545 AA.
AC P28769;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE T-COMPLEX PROTEIN 1, ALPHA SUBUNIT (TCP-1-ALPHA) (CCT-ALPHA).
GN CCT1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC eubryophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC Arabidopsis.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 93138412.
RA MORI M., MURATA K., KUBOTA H., YAMAMOTO A., MATSUSHIRO A.,
RA MORITA T.;
RT "Cloning of a cDNA encoding the Tcp-1 (t complex polypeptide 1)
RT homologue of Arabidopsis thaliana."
RL Gene 122:381-382(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLOMBIA;
RA MORI M., KUNO N., MURATA K., KUBOTA H., FURUYA M., MATSUSHIRO A.,
RA MORITA T.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MOLECULAR CHAPERONE: ASSIST THE FOLDING OF PROTEINS UPON
CC APF HYDROLYSIS. KNOWN TO PLAY A ROLE, IN VITRO, IN THE FOLDING OF
CC ACTIN AND TUBULIN.
CC -!- SUBUNIT: HETERO-OLIGOMERIC COMPLEX OF ABOUT 850 TO 900 KD THAT
CC FORMS TWO STACKED RINGS, 12 TO 16 NM IN DIAMETER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE TCP-1 CHAPERONIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D11351; BAA01955.1; -.
CC DR EMBL; D11352; BAA21772.1; -.
CC DR PIR; JN0448; JN0448.
CC DR HSSP; P48424; IASX.
CC DR PROSITE; PS00750; TCP1_1; 1.
CC DR PROSITE; PS00751; TCP1_2; 1.
CC DR PROSITE; PS00995; TCP1_3; 1.
CC DR PFAM; PF00118; cpn60_TCP1; 1.
CC KW Chaperone; ATP-binding; Multigene family.
CC SEQUENCE 545 AA; 59229 MW; 8E52A94B CRC32;

OY 1 YRLAIR 6
    ||||:
Db 124 YRLAMR 129

Query Match 89.7%; Score 26; DB 1; Length 545;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 7
ID 3DHQ_BACSU STANDARD; PRT; 148 AA.
AC P54517;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PUTATIVE CATABOLIC 3-DEHYDROQUINATE DEHYDRATASE (EC 4.2.1.10) (3-
DE DEHYDROQUINASE).
GN YQHS.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-168 / JH642;
RA KOBAYASHI Y., MIZUNO M., MASUDA S., TAKEMARU K., HOSONO S.,
RA SATO T., TAKEUCHI M.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 3-DEHYDROQUINATE -> 3-DEHYDROSHIKIMATE + H(2)O.
CC -1- PATHWAY: SECOND STEP OF THE QUINIC ACID CATABOLIC PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE TYPE-II 3-DEHYDROQUINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; D84432; BAA12556.1; -
DR EMBL; Z99116; CAB14378.1; -
DR SUBTILIST; BG11707; YQHS.
DR PROSITE; PS01029; DEHYDROQUINASE_II; 1.
DR PFAM; PF01220; DHQINASE_II; 1.
KW Hypothetical protein; Quinate metabolism; Lyase.
SQ SEQUENCE 148 AA; 16431 MW; B2D0F289 CRC32;

Query Match 86.2%; Score 25; DB 1; Length 148;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
Db 132 YKLAVR 137
|:|:|

RESULT 8
FLIZ_ECOLI
ID FLIZ_ECOLI STANDARD; PRT; 183 AA.
AC P52627; P76317;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE FLIZ PROTEIN.
GN FLIZ.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-RP437;
RX MEDLINE; 96125225.
RA MYTELKA D.S., CHAMBERLIN M.J.;
RT "Escherichia coli flizy operon."
RL J. Bacteriol. 178:24-34(1996).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;

Query Match 86.2%; Score 25; DB 1; Length 183;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
Db 152 YRLAIR 157
|:|:|

RESULT 9
COOF_RHURU
ID COOF_RHURU STANDARD; PRT; 190 AA.
AC P31894;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE IRON-SULFUR PROTEIN.
GN COOF.
OS Rhodospirillum rubrum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
OC Rhodospirillum.
[1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-URL;
RX MEDLINE; 92355502.

```

RA KERBY R.L., HONG S.S., ENSIGN S.A., COPPOC L.J., LUDDEN P.W.,
 RA ROBERTS G.P.;
 RA "Genetic and physiological characterization of the Rhodospirillum
 RT rubrum carbon monoxide dehydrogenase system.";
 RL J. Bacteriol. 174:5284-5294(1992).
 CC -!- FUNCTION: COB MAY OXYDATE CARBON MONOXIDE COUPLED, VIA COOF, TO
 CC REDUCTION OF HYDROGEN CATION BY AN HYDROGENASE (POSSIBLY COOH).
 CC COOF IS REQUIRED IN STOICHIOMETRIC AMOUNTS IN VITRO FOR ANCHORING
 CC COB TO THE MEMBRANE AS WELL AS FOR CONVEYING THE ELECTRONS TO
 CC THE HYDROGENASE.
 CC -!- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
 CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U65510; AAC45122.1; -
 DR PIR: B42957; B42957.
 DR HSSP: P00208; 1BLG.
 DR PROSITE: PS00198; 4FE4S_FERREDOXIN; 1.
 DR PFAM: PF00037; fer4; 2.
 KW Iron-sulfur; Electron transport; 4Fe-4S.
 FT METAL 17 17 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
 FT METAL 20 20 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
 FT METAL 23 23 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
 FT METAL 27 27 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
 FT METAL 65 65 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 68 68 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 73 73 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 77 77 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 96 96 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
 FT METAL 99 99 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
 FT METAL 102 102 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
 FT METAL 106 106 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
 FT METAL 133 133 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
 FT METAL 136 136 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
 FT METAL 150 150 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
 FT METAL 154 154 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
 SQ SEQUENCE 190 AA; 20567 MW; 6C6B83ED CRC32;
 Query Match 86.2%; Score 25; DB 1; Length 190;
 Best Local Similarity 66.7%; Pred. No. 24;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YRLAIR 6
 Db 167 YRLAIR 172
 RESULT 10
 COAT_CTV36
 ID COAT_CTV36 STANDARD; PRT; 223 AA.
 AC Q00686;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE COAT PROTEIN.
 OS Citrus tristezza virus (isolate T36) (CTV).
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
 CC Closterovirus.
 CC [1]
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 27-38.
 RX MEDLINE: 91237334.
 RA SEKIYA M.E., LAWRENCE S.D., MCCAFFERY M., CLINE K.;
 RT "Molecular cloning and nucleotide sequencing of the coat protein gene
 RT of citrus tristezza virus.";
 RL J. Gen. Virol. 72:1013-1020(1991).

[2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE: 94160579.
 RA PAPPU H.R., KARASEV A.V., ANDERSON E.J., PAPPU S.S., HILF M.E.,
 RA FERRIS V., ECKLOFF R.M.G., MCCAFFERY M., BOYKO V., GOMDA S.,
 RA DOLJA V.V., KOONIN E.V.;
 RT "Nucleotide sequence and organization of eight 3' open reading frames
 RT of the citrus tristezza closterovirus genome.";
 RL Virology 199:35-46(1994).
 CC -!- PTM: CONSISTS OF AT LEAST TWO SIZE VARIANTS, CP1 AND CP2, WHICH
 CC RESULT OF POST-TRANSLATIONAL PROTEOLYSIS AT SITES APPROXIMATIVELY
 CC 12 TO 15 AND 26 AA FROM THE N-TERMINUS RESPECTIVELY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M76485; AAA42924.1; -
 DR EMBL: U16304; AAC59630.1; -
 KW Coat protein.
 SQ SEQUENCE 223 AA; 24909 MW; F3AA6BB6 CRC32;
 Query Match 86.2%; Score 25; DB 1; Length 223;
 Best Local Similarity 66.7%; Pred. No. 28;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YRLAIR 6
 Db 85 YRLAVK 90
 RESULT 11
 FRDC_WOLSU
 ID FRDC_WOLSU STANDARD; PRT; 256 AA.
 AC P17413.
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-JUL-1998 (Rel. 36, Last annotation update)
 DE FUMARATE REDUCTASE CYTOCHROME B SUBUNIT.
 GN FRDC.
 OS Wolinella succinogenes.
 CC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 CC Wolinella.
 CC [1]
 RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE: 90355847.
 RA KOETNER C., LAUTERBACH F., TRIPTER D., UNDEN G., KROEGER A.;
 RT "Wolinella succinogenes fumarate reductase contains a dihaem
 RT cytochrome b.";
 RL Mol. Microbiol. 4:855-860(1990).
 RN [2]
 RN SEQUENCE FROM N.A.
 RA SIMON J., GROSS R., RINGEL M., SCHMIDT E., KROEGER A.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: DI-HEME CYTOCHROME OF THE FUMARATE REDUCTASE COMPLEX.
 CC -!- SUBUNIT: PART OF AN ENZYME COMPLEX CONTAINING THREE SUBUNITS:
 CC A FLAVOPROTEIN, AN IRON-SULFUR PROTEIN AND A CYTOCHROME B.
 CC -!- SIMILARITY: TO CYTOCHROME B-558 FROM B.SUBTILIS SUCCINATE
 CC DEHYDROGENASE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: X51509; CAA35874.1; --
 DR EMBL: AJ000662; CAA04213.1; --
 DR PIR: S10164; S10164.
 KW Tricarboxylic acid cycle; Transmembrane; Electron transport; Heme.
 FT TRANSMEM 32 54 POTENTIAL.
 FT TRANSMEM 77 98 POTENTIAL.
 FT TRANSMEM 123 149 POTENTIAL.
 FT TRANSMEM 171 192 POTENTIAL.
 FT TRANSMEM 214 232 POTENTIAL.
 FT BINDING 44 44 HEME (POTENTIAL).
 FT BINDING 93 93 HEME (POTENTIAL).
 FT BINDING 143 143 HEME (POTENTIAL).
 FT BINDING 182 182 HEME (POTENTIAL).
 SQ SEQUENCE 256 AA; 29723 MW; C3AB8928 CRC32;
 Query Match 86.2%; Score 25; DB 1; Length 256;
 Best Local Similarity 66.7%; Pred. No. 32;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YRLAIR 6
 Db 188 YRLAVK 193
 RESULT 12
 FRET_YEAST
 ID FRET_YEAST STANDARD; PRT; 629 AA.
 AC Q12333;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE FERRIC REDUCTASE TRANSMEMBRANE COMPONENT 7 PRECURSOR.
 GN FRET OR YOL152W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;
 RX MEDLINE; 96132030.
 RA CASAYAYOR A., ALDEA M., CASAS C., HERRERO E., GAMO F.J.,
 RA LAFUENTE M.J., GANCEDO C., ARINO J.;
 RT "DNA sequence analysis of a 13 kbp fragment of the left arm of yeast
 chromosome XV containing seven new open reading frames.";
 RL yeast 11:1281-1288(1995).
 CC -1- COFACTOR: FAD (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE FRE / CYBB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z48239; CAA88276.1; --
 DR EMBL: Z74894; CAA99174.1; --
 KW Electron transport; Transmembrane; Iron transport; FAD; NAD;
 KW Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 629
 FT CHAIN ? 629
 FT NP_BIND 369 375 FAD (POTENTIAL).
 FT TRANSMEM 46 66 POTENTIAL.
 FT TRANSMEM 108 128 POTENTIAL.
 FT TRANSMEM 168 188 POTENTIAL.
 FT TRANSMEM 195 215 POTENTIAL.
 FT TRANSMEM 238 258 POTENTIAL.
 FT TRANSMEM 266 286 POTENTIAL.

FT TRANSMEM 293 313 POTENTIAL.
 FT TRANSMEM 422 442 POTENTIAL.
 FT CARBOHYD 330 330 POTENTIAL.
 FT CARBOHYD 541 541 POTENTIAL.
 SQ SEQUENCE 629 AA; 71996 MW; AC37EFC2 CRC32;
 Query Match 86.2%; Score 25; DB 1; Length 629;
 Best Local Similarity 66.7%; Pred. No. 84;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YRLAIR 6
 Db 309 YRLAVK 314
 RESULT 13
 AROL_EMENI
 ID AROL_EMENI STANDARD; PRT; 1586 AA.
 AC P07547;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 15-DEC-1999 (Rel. 39, Last sequence update)
 DT 15-DEC-1999 (Rel. 39, Last annotation update)
 DE PENTAFUNCTIONAL AROM POLYPEPTIDE [INCLUDES: 3-DEHYDROQUINATE SYNTHASE
 (EC 4.6.1.3); 3-DEHYDROQUINATE DEHYDRATASE (EC 4.2.1.10)
 (3-DEHYDROQUINASE); SHIKIMATE 5-DEHYDROGENASE (EC 1.1.1.25); SHIKIMATE
 KINASE (EC 2.7.1.71); EPSF SYNTHASE (EC 2.5.1.19)].
 GN AROMA OR AROM.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Euascomycetes; Plecomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RL53;
 RX MEDLINE; 86176723.
 RA CHARLES I.G., KEYTE J.W., BRAMMAR W.J., SMITH M., HAWKINS A.R.;
 RT "The isolation and nucleotide sequence of the complex AROM locus of
 Aspergillus nidulans.";
 RL Nucleic Acids Res. 14:2201-2213(1986).
 RN [2]
 RP SEQUENCE OF 844-1474 FROM N.A.
 RX MEDLINE; 86067221.
 RA CHARLES I.G., KEYTE J.W., BRAMMAR W.J., HAWKINS A.R.;
 RT "Nucleotide sequence encoding the biosynthetic dehydroquinase
 function of the penta-functional arom locus of Aspergillus
 nidulans.";
 RL Nucleic Acids Res. 13:8119-8128(1985).
 RN [3]
 RP REVISIONS TO THE C-TERMINUS.
 RC STRAIN=RL53;
 RA HAWKINS A.R.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE AROM POLYPEPTIDE CATALYZES 5 CONSECUTIVE ENZYMATIC
 CC REACTIONS IN PRECHORISMATE POLYAROMATIC AMINO ACID BIOSYNTHESIS.
 CC -1- CATALYTIC ACTIVITY: 7-PHOSPHO-3-DEOXY-ARABINO-HEPTULOSONATE -
 CC 3-DEHYDROQUINATE + ORTHOPHOSPHATE.
 CC -1- CATALYTIC ACTIVITY: 3-DEHYDROQUINATE -> 3-DEHYDROSHIKIMATE + H(2)O.
 CC -1- CATALYTIC ACTIVITY: SHIKIMATE + NADP(+) -> 5-DEHYDROSHIKIMATE +
 CC NADPH.
 CC -1- CATALYTIC ACTIVITY: ADP + SHIKIMATE -> ADP + SHIKIMATE 3-PHOSPHATE.
 CC -1- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHAKIMATE ->
 CC ORTHOPHOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHAKIMATE.
 CC -1- PATHWAY: SECOND TO SIXTH STEP IN THE BIOSYNTHESIS FROM CHORISMATE
 CC OF THE AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```
DR EMBL; X05204; CAA28836.1; -.
DR PIR; A24042; BVASAL.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
DR PROSITE; PS01028; DEHYDROQUINASE_I; 1.
DR PROSITE; PS01128; SHIKIMATE_KINASE; 1.
DR PFAM; PF00275; EPSP_synthase; 1.
DR PFAM; PF01202; SKI; 1.
DR PFAM; PF01487; DHquinase_I; 1.
DR PFAM; PF01488; Shikimate_DH; 1.
KW Aromatic amino acid biosynthesis; Multifunctional enzyme;
FT Oxidoreductase; Lyase; Transferase; Kinase; NADP; ATP-binding.
FT DOMAIN 1 384 3-DEHYDROQUINATE SYNTHASE.
FT DOMAIN 397 843 EPSP SYNTHASE.
FT DOMAIN 863 1056 SHIKIMATE KINASE.
FT DOMAIN 1057 1277 3-DEHYDROQUINASE.
FT DOMAIN 1290 1586 SHIKIMATE DEHYDROGENASE.
FT ACT_SITE 825 825 POTENTIAL.
FT NP_BIND 871 878 ATP (BY SIMILARITY).
FT ACT_SITE 1180 1180 BY SIMILARITY.
FT ACT_SITE 1208 1208 FORMS A SHIFT-BASE INTERMEDIATE
                          (BY SIMILARITY).
SQ SEQUENCE 1586 AA; 173186 MW; 279EFF15 CRC32;

Query Match      86.2%; Score 25; DB 1; Length 1586;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIR 6
   |||||
DB 1141 YRLAFR 1146

RESULT 14
ID YF65_SYNY3 STANDARD; PRT; 113 AA.
AC P74596;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 12.5 KD PROTEIN SLR1565.
GN SLR1565.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97061201.
RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K.,
RA OKUMURA S., SHIMPO S., TAKEUCHI C., WADA T., WATANABE A.,
RA YAWADA M., YASUDA M., TABATA S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1995).
CC -1- SIMILARITY: BELONGS TO THE HESB/YADB/YEHP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90916; BAA18704.1; -.
DR PROSITE; PS01152; HESB; 1.
DR PFAM; PF01521; Hesb-like; 1.
KW Hypothetical protein.
SQ SEQUENCE 113 AA; 12520 MW; 927AED62 CRC32;

Query Match      82.8%; Score 24; DB 1; Length 113;
Best Local Similarity 66.7%; Pred. No. 25;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
   :|||:|
DB 26 FRLAVR 31

RESULT 15
ID YG2E_YEAST STANDARD; PRT; 123 AA.
AC P53247;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL 14.1 KD PROTEIN IN UPF3-SMD1 INTERGENIC REGION.
GN YGR073C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RA WEDLER H., SCHARFE M., WEDLER E., WAMBUIT R.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z72858; CAA97075.1; -.
DR Hypothetical protein; Transmembrane.
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 103 123 POTENTIAL.
SQ SEQUENCE 123 AA; 14066 MW; 9EDC70DB CRC32;

Query Match      82.8%; Score 24; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAI 5
   |||||
DB 27 YRLAI 31

Search completed: February 8, 2000, 00:59:49
Job time: 3778 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 13:17:33 ; Search time 209.03 Seconds

(without alignments)
1.990 Million cell updates/sec

Title: US-08-653-294-7

Perfect score: 29

Sequence: 1 YRLAIR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

SPTREMBL_12:.*
1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mhc:.*
8: sp_organelle:.*
9: sp_phase:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-----------|--------------------|
| 1 | 29 | 100.0 | 261 | 2 Q92671 | Q92671 Zymomonas m |
| 2 | 29 | 100.0 | 556 | 3 Q94501 | Q94501 schizosacch |
| 3 | 28 | 96.6 | 343 | 3 P72467 | P72467 streptomyc |
| 4 | 28 | 96.6 | 350 | 5 O45803 | O45803 caenorhabd |
| 5 | 28 | 96.6 | 545 | 12 P89475 | P89475 herpes simp |
| 6 | 28 | 96.6 | 579 | 2 Q9WZ01 | Q9WZ01 thermotoga |
| 7 | 27 | 93.1 | 185 | 2 Q55192 | Q55192 synecocyst |
| 8 | 27 | 93.1 | 283 | 2 Q9ZA48 | Q9ZA48 streptomyc |
| 9 | 27 | 93.1 | 309 | 2 O86347 | O86347 mycobacteri |
| 10 | 27 | 93.1 | 404 | 2 O83460 | O83460 treponema p |
| 11 | 27 | 93.1 | 845 | 5 O01314 | O01314 caenorhabd |
| 12 | 27 | 93.1 | 1280 | 13 Q90933 | Q90933 gallus gall |
| 13 | 26 | 89.7 | 96 | 1 O29637 | O29637 archaeglob |
| 14 | 26 | 89.7 | 255 | 2 Q9ZMN9 | Q9ZMN9 helicobacte |
| 15 | 26 | 89.7 | 291 | 1 Q9V560 | Q9V560 aeropyrum p |
| 16 | 26 | 89.7 | 362 | 5 Q22521 | Q22521 caenorhabd |
| 17 | 26 | 89.7 | 384 | 2 P97084 | P97084 salmonella |
| 18 | 26 | 89.7 | 809 | 2 Q9ZGH5 | Q9ZGH5 streptomyc |
| 19 | 26 | 89.7 | 883 | 5 Q07995 | Q07995 chironomus |
| 20 | 25 | 86.2 | 109 | 12 O11376 | O11376 molluscum c |

| | | | | | |
|----|----|------|-----|-----------|--------------------|
| 21 | 25 | 86.2 | 110 | 12 Q98316 | Q98316 molluscum c |
| 22 | 25 | 86.2 | 154 | 12 Q89217 | Q89217 citrus tris |
| 23 | 25 | 86.2 | 154 | 12 Q89225 | Q89225 citrus tris |
| 24 | 25 | 86.2 | 154 | 12 Q89227 | Q89227 citrus tris |
| 25 | 25 | 86.2 | 154 | 12 Q89229 | Q89229 citrus tris |
| 26 | 25 | 86.2 | 154 | 12 Q91468 | Q91468 citrus tris |
| 27 | 25 | 86.2 | 154 | 12 Q92005 | Q92005 citrus tris |
| 28 | 25 | 86.2 | 154 | 12 Q92004 | Q92004 citrus tris |
| 29 | 25 | 86.2 | 154 | 12 Q92003 | Q92003 citrus tris |
| 30 | 25 | 86.2 | 154 | 12 Q92002 | Q92002 citrus tris |
| 31 | 25 | 86.2 | 154 | 12 Q92001 | Q92001 citrus tris |
| 32 | 25 | 86.2 | 154 | 12 Q92000 | Q92000 citrus tris |
| 33 | 25 | 86.2 | 154 | 12 Q9W858 | Q9W858 citrus tris |
| 34 | 25 | 86.2 | 162 | 12 Q66251 | Q66251 citrus tris |
| 35 | 25 | 86.2 | 192 | 3 O13610 | O13610 schizosacch |
| 36 | 25 | 86.2 | 223 | 12 O10474 | O10474 citrus tris |
| 37 | 25 | 86.2 | 223 | 12 Q06192 | Q06192 citrus tris |
| 38 | 25 | 86.2 | 223 | 12 P89948 | P89948 citrus tris |
| 39 | 25 | 86.2 | 223 | 12 Q9WID9 | Q9WID9 citrus tris |
| 40 | 25 | 86.2 | 368 | 2 Q9X0U3 | Q9X0U3 thermotoga |
| 41 | 25 | 86.2 | 380 | 2 O05875 | O05875 mycobacteri |
| 42 | 25 | 86.2 | 449 | 5 O62027 | O62027 caenorhabd |
| 43 | 25 | 86.2 | 463 | 1 Q9YAG3 | Q9YAG3 aeropyrum p |
| 44 | 25 | 86.2 | 545 | 10 O64767 | O64767 arabidopsis |
| 45 | 25 | 86.2 | 744 | 2 O33087 | O33087 mycobacteri |

ALIGNMENTS

RESULT 1

ID Q92671 PRELIMINARY; PRT; 261 AA.
AC Q92671
DT 01-MAY-1999 (TREMREL. 10, Created)
DT 01-MAY-1999 (TREMREL. 10, Last sequence update)
DT 01-NOV-1999 (TREMREL. 12, Last annotation update)
DE THYMIDYLATE SYNTHETASE.
GN THYA.
OS Zymomonas mobilis.
OC Bacteria; Proteobacteria; alpha subdivision; Zymomonas group;
OC Zymomonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZM4.
RA UM H.W., KANG H.S.;
RT "Sequence analysis of 43A9 fosmid clone of Zymomonas mobilis ZM4.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + DUMP =
CC DIHYDROFOLATE + DTMP.
CC -1- PATHWAY: DEOXYRIBONUCLEOTIDE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC EMBL: AF102543; AADI9406.1; -.
DR HSP; P00470; IAN5.
DR PROSITE; PS00091; THYMIDYLATE SYNTHASE; 1.
KW Transferase; Methyltransferase; Nucleotide biosynthesis.
FT ACT_SITE 143 143 BY SIMILARITY.
SQ SEQUENCE 261 AA; 29755 MW; A2D870D0 CRC32;

Query Match 100.0%; Score 29; DB 2; Length 261;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
|||||
DB 21 YRLAIR 26

RESULT 2

ID Q94501 PRELIMINARY; PRT; 556 AA.
AC Q94501;

DT 01-MAY-1999 (TRENBLrel. 10, Last Created)
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE T-COMPLEX PROTEIN 1, ALPHA SUBUNIT HOMOLOG, CHAPERONIN FAMILY.
 GN SPBC12D12.03.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972H-;
 RA LYNE M., RAJANDREAM M.A., BARRELL B.G., XIANG Z., AVES S;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL035085; CAA22677.1; -;
 DR HSP; P48425; IAGE.
 DR PROSITE; PS00750; TCPI_1; 1.
 DR PROSITE; PS00751; TCPI_2; 1.
 DR PROSITE; PS00995; TCPI_3; 1.
 SQ SEQUENCE 556 AA; 60047 MW; 27F285FA CRC32;

 Query Match 100.0%; Score 29; DB 3; Length 556;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 YRLAIR 6
 DB 126 YRLAIR 131
 |||||

 RESULT 3
 ID P72467 PRELIMINARY; PRT; 343 AA.
 AC P72467;
 DT 01-FEB-1997 (TRENBLrel. 02, Created)
 DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
 DE ORF4.
 OS Streptomyces lividans.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-66;
 RX MEDLINE; 97286540.
 RA BETZLER M., TIOLKA I., SCHREMPF H.;
 RT "Amplification of a Streptomyces lividans 4.3 kb DNA element causes overproduction of a novel hypha- and vesicle-associated protein.";
 RL Microbiology 143:1243-1252(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-66;
 RX MEDLINE; 88121703.
 RA FORSMAN M., JAURIN B.;
 RT "Chromogenic identification of promoters in Streptomyces lividans by using an ampC beta-lactamase promoter-probe vector.";
 RL Mol. Gen. Genet. 210:23-32(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-66;
 RX MEDLINE; 95091658.
 RA LABESSE G., VIDAL-GROS A., CHOMILIER J., GAUDRY M., MORNON J.;
 RT "Structural comparisons lead to the definition of a new superfamily of NAD(P)(H)-accepting oxidoreductases: the single-domain reductases/epimerases/dehydrogenases (the 'RED' family).";
 RL Biochem. J. 304:95-99(1994).
 DR EMBL; 270724; CAA94732.1; -;
 SQ SEQUENCE 343 AA; 36846 MW; 90612F52 CRC32;

 Query Match 96.6%; Score 28; DB 2; Length 343;
 Best Local Similarity 83.3%; Pred. No. 33;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 YRLAIR 6
 DB 227 YRLAIR 232
 |||||

 RESULT 4
 ID 045803 PRELIMINARY; PRT; 350 AA.
 AC 045803;
 DT 01-JUN-1998 (TRENBLrel. 06, Created)
 DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
 DE T2D5.1 PROTEIN.
 GN T2D5.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA LLOYD C., WILKINSON J.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 94150718.
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A., BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L., CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P., LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M., PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R., SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R., WAYSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";
 RL Nature 368:32-38(1994).
 DR EMBL; 282051; CAB04815.1; -;
 DR PFAM; PF01461; 7tm_4; 1.
 SQ SEQUENCE 350 AA; 40291 MW; A9117B4D CRC32;

 Query Match 96.6%; Score 28; DB 5; Length 350;
 Best Local Similarity 83.3%; Pred. No. 34;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 YRLAIR 6
 DB 307 YRLAIR 312
 |||||

 RESULT 5
 ID P89475 PRELIMINARY; PRT; 545 AA.
 AC P89475;
 DT 01-MAY-1997 (TRENBLrel. 03, Created)
 DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
 DE VIRION GLYCOPROTEIN E.
 GN US8.
 OS Herpes simplex virus (type 2).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HG52;
 RX MEDLINE; 87111457.
 RA MCGEOCH D.J., MOSS H.W., MCNAB D., FRAME M.C.;
 RT "DNA sequence and genetic content of the Herpes simplex virus type 2 genome: identification of the gene encoding glycoprotein G, and evolutionary

RT comparisons.";
 RL J. Gen. Virol. 68:19-38(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HG52;
 RX MEDLINE; 90278430.
 RA EVRETT R., FENWICK M.;
 RT "Comparative DNA sequence analysis of the host shutoff genes of
 RT different strains of herpes simplex virus: type 2 strain HG52 encodes
 RT a truncated UL41 product.";
 RL J. Gen. Virol. 71:1387-1390(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HG52;
 RX MEDLINE; 92113549.
 RA MCGEOCH D.J., CUNNINGHAM C., MCINTYRE G., DOLAN A.;
 RT "Comparative sequence analysis of the long repeat regions and
 RT adjoining parts of the long unique regions in the genomes of herpes
 RT simplex viruses types 1 and 2.";
 RL J. Gen. Virol. 72:3057-3075(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HG52;
 RX MEDLINE; 92356101.
 RA BARNETT B.C., DOLAN A., TELFORD E.A.R., DAVISON A.J., MCGEOCH D.J.;
 RT "A novel herpes simplex virus gene (UL49A) encodes a putative membrane
 RT protein with counterparts in other herpesviruses.";
 RL J. Gen. Virol. 73:2167-2171(1992).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HG52;
 RA DOLAN A.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z86099; CAB06715.1; -;
 SQ SEQUENCE 545 AA; 59309 MW; 99AB4CDD CRC32;

Query Match 96.6%; Score 28; DB 12; Length 545;
 Best Local Similarity 83.3%; Pred. No. 54;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YRLAIR 6
 ||||:|
 Db 299 YRLAVR 304

RESULT 6
 Q9WZ01
 ID Q9WZ01 PRELIMINARY; PRT; 579 AA.
 AC Q9WZ01;
 DT 01-NOV-1999 (TRENBLREL. 12, Created)
 DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)
 DT 01-NOV-1999 (TRENBLREL. 12, Last annotation update)
 DE OLIGOPEPTIDE ABC TRANSPORTER, PERIPLASMIC OLIGOPEPTIDE-BINDING
 DE PROTEIN.
 GN TW0531.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 99287316.
 RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,
 RA HART D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,
 RA MCDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,
 RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,
 RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,
 RA SMITH H.O., VENTER J.C., FRASER C.M.;
 RT "Evidence for lateral gene transfer between Archaea and bacteria from
 RT genome sequence of Thermotoga maritima.";
 RL Nature 399:323-329(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,

RA HART D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,
 RA MCDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,
 RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,
 RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,
 RA SMITH H.O., VENTER J.C., FRASER C.M.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE001728; AAD35616.1; -;
 SQ SEQUENCE 579 AA; 66508 MW; BF5C4AB3 CRC32;

Query Match 96.6%; Score 28; DB 2; Length 579;
 Best Local Similarity 83.3%; Pred. No. 57;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YRLAIR 6
 ||||:|
 Db 168 YRLAVR 173

RESULT 7
 Q55192
 ID Q55192 PRELIMINARY; PRT; 185 AA.
 AC Q55192;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-JAN-1999 (TRENBLREL. 09, Last annotation update)
 DE HYPOTHETICAL 20.8 KD PROTEIN.
 OS Synecocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PCC6803;
 RA TABATA S.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PCC6803;
 RX MEDLINE; 96127529.
 RA KANEKO T., TANAKA A., SATO S., KOTANI H., SAZUKA T., MIYAJIMA N.,
 RA SUGIURA M., TABATA S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synecocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
 RT region from map positions 64% to 92% of the genome.";
 RL DNA Res. 2:153-166(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PCC6803;
 RX MEDLINE; 97061201.
 RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
 RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
 RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,
 RA SHIMPO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,
 RA TABATA S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL; D64001; BAA10334.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 185 AA; 20830 MW; 365A078D CRC32;

Query Match 93.1%; Score 27; DB 2; Length 185;
 Best Local Similarity 83.3%; Pred. No. 31;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YRLAIR 6
 ||||:|
 Db 35 YRLAIR 40

RESULT 8
 Q9ZA48

```

ID Q9ZA48 PRELIMINARY; PRT; 283 AA.
AC
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
DE PATHWAY-SPECIFIC TRANSCRIPTIONAL ACTIVATOR.
GN GRA-ORF9.
OS Streptomyces violaceoruber.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TU22;
RX MEDLINE; 99051446.
RA ICHINOSE K., BEDFORD D.J., TORNUS D., BECHTHOLD A., BIBB M.J.,
RA REVILL W.P., FLOSS H.G., HOPWOOD D.A.;
RT "The granaticin biosynthetic gene cluster of Streptomyces
RT violaceoruber Tu22: sequence analysis and expression in a heterologous
RT host.";
RL Chem. Biol. 5:647-659(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TU22;
RX MEDLINE; 90060034.
RA SHERMAN D.H., MALPARTIDA F., BIBB M.J., KIESER H.M., BIBB M.J.,
RA HOPWOOD D.A.;
RT "Structure and deduced function of the granaticin-producing polyketide
RT synthase gene cluster of Streptomyces violaceoruber Tu22.";
RL EMBO J. 8:2717-2725(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-TU22;
RX MEDLINE; 96027933.
RA BECHTHOLD A., SOHNG J.K., SMITH T.M., CHU X., FLOSS H.G.;
RT "Identification of Streptomyces violaceoruber Tu22 genes involved in
RT the biosynthesis of granaticin.";
RL Mol. Gen. Genet. 248:610-620(1995).
DR EMBL; AJ011500; CAA09630.1; -.
SQ SEQUENCE 283 AA; 30486 MW; C24226DC CRC32;

Query Match 93.1%; Score 27; DB 2; Length 283;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
DB 109 YRLAIR 114

RESULT 9
O86347 PRELIMINARY; PRT; 309 AA.
AC
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE HYPOTHETICAL 33.5 KD PROTEIN.
GN RV2776C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE; 98295987.
RA COLE S.T., BROSCHE R., PARKHILL J., GARNIER T., CHURCHER C., HARRIS D.,
RA GORDON S.V., EIGLMEIER K., GAS S., BARRY III C.E., TEKAIA F.,
RA BADCOCK K., BASHAM D., BROWN D., CHILLINGWORTH T., CONNOR R.,
RA DAVIES R., DEVLIN K., FELTWELL T., GENTLES S., HAMLIN N., HOLROYD S.,
RA HORNSBY T., JAGELS K., KROGH A., MCLEAN J., MOULE S., MURPHY L.,
RA OLIVER S., OSBORNE J., QUAIL M.A., RAJANDREAM M.A., ROGERS J.,
RA RUTTER S., SEEGER K., SKELTON S., SQUARES S., SOARES R., SULSTON J.E.,

```

```

RA TAYLOR K., WHITEHEAD S., BARRELL B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RN Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA PARKHILL J.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL008967; CAA15591.1; -.
DR HSSP; P33164; 2PIA.
DR PROSITE; PS00197; 2FE2S-FERREDOXIN; 1.
DR PFAM; PF00111; fer2; 1.
DR PFAM; PF00175; oxidored_fad; 1.
KW Hypothetical protein; Iron-sulfur.
SQ SEQUENCE 309 AA; 33517 MW; B152B590 CRC32;

Query Match 93.1%; Score 27; DB 2; Length 309;
Best Local Similarity 83.3%; Pred. No. 53;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
DB 65 YRLAIR 70

RESULT 10
O83460 PRELIMINARY; PRT; 404 AA.
AC
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE GCPE PROTEIN (GCPE).
GN TP0446.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98332770.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.,
RA DODSON R., GWINN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.,
RA SODERGREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.,
RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTTERBACK T.,
RA McDONALD L., ARTIACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.,
RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J., SMITH H.O.,
RA VENTER J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT Spirochete";
RL Science 281:375-388(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.,
RA DODSON R., GWINN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.,
RA SODERGREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.,
RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTTERBACK T.,
RA McDONALD L., ARTIACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.,
RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J., SMITH H.O.,
RA VENTER J.C.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE001221; AAC65433.1; -.
DR TIGR; TP0446; -.
SQ SEQUENCE 404 AA; 43733 MW; DA02EFA40 CRC32;

Query Match 93.1%; Score 27; DB 2; Length 404;
Best Local Similarity 83.3%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIR 6
DB 121 YRLAIR 126

```

Query Match 89.7%; Score 26; DB 1; Length 96;

Best Local Similarity 83.3%; Pred. No. 28;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YRLAIR 6
|||||
Db 51 YRLAIK 56

RESULT 14

Q9ZMN9 PRELIMINARY; PRT; 255 AA.
ID Q9ZMN9;
AC Q9ZMN9;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE FUMARATE REDUCTASE.
GN FRDC.
OS Helicobacter pylori J99.
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-J99;
RX MEDLINE; 99120557.
RA ALM R.A., LING L.-S.L., MOIR D.T., KING B.L., BROWN E.D., DOIG P.C.,
RA SMITH D.R., NOONAN B., GUILD B.C., DEJONGE B.L., CARMEL G.,
RA TUNMINO P.J., CARUSO A., URIA-NICKELSEN M., MILLS D.M., IVES C.,
RA GIBSON R., MERBERG D., MILLS S.D., JIANG Q., TAYLOR D.E., VOVIS G.F.,
RA TRUST T.J.;
RT "Genomic-sequence comparison of two unrelated isolates of the human
gastric pathogen Helicobacter pylori";
RL Nature 397:176-180(1999).
DR EMBL; AE001456; AAD05763.1; -.
SQ SEQUENCE 255 AA; 28872 MW; 04B2EB82 CRC32;

Query Match 89.7%; Score 26; DB 2; Length 255;

Best Local Similarity 83.3%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YRLAIR 6
|||||
Db 188 YRLAIK 193

RESULT 15

Q9YE60 PRELIMINARY; PRT; 291 AA.
ID Q9YE60;
AC Q9YE60;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE 291AA LONG HYPOTHETICAL PROTEIN.
GN APE0710.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Aeropyrum.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K1;
RX MEDLINE; 99310339.
RA KAWARABAYASI Y., HINO Y., HORIKAWA H., YAMAZAKI S., HAIKAWA Y.,
RA JIN-NO K., TAKAHASHI M., SEKINE M., BABA S., ANKAI A., KOSUGI H.,
RA HOSOIYAMA A., FUKUI S., NAGAI Y., NISHIJIMA K., NAKAZAWA H.,
RA TAKAMIZAKI M., MASUDA S., FUNAHASHI T., TANAKA T., KUDOH Y.,
RA YAMAZAKI J., KUSHIDA N., OGUCHI A., AOKI K., KUBOTA K., NAKAMURA Y.,
RA NOMURA N., SAKO Y., KIKUCHI H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
crenarchaeon, Aeropyrum pernix K1";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000080; BAA79686.1; -.
SQ SEQUENCE 291 AA; 31756 MW; 1D26675E CRC32;

Query Match 89.7%; Score 26; DB 1; Length 291;
Best Local Similarity 83.3%; Pred. No. 88;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YRLAIR 6
|||||
Db 66 YRLAMR 71

Search completed: February 8, 2000, 13:17:35
Job time: 32484 sec

OM of: US-08-653-294-7 to: GenEmbl.* out_format : pfs

Date: Feb 8, 2000 4:37 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL-frame+ pzn.model -DEV-xlp
-O/cgml_1/USPTO.spool/US08653294/runat_04022000_160701_15779/app_query.fasta.1
-DB-GenEmbl -GFMF-fastap -SUFFIX-ige -GAPOP-12.000 -GAPEXT-4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -OGAPOP-4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -EGAPOP-6.000
-XGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP-6.000
-DELEXT=7.000 -START=1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -ALIGN=15 -MODE-LOCAL
-OUTFMT=pfs -NORM-ext -MINLEN=0 -MAXLEN=1000000 -USER-US08653294
-NCPU=6 -ICPU=3 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-7

Query length: 6

Database: GenEmbl.*

Database sequences: 821193

Database length: -1518192014

Search time (sec): 11370.480000

score_list:

| Sequence | Strd Orig | ZScore | Escore | Len | Documentation |
|------------------|-----------|--------|--------|---------|-----------------------------------|
| gb_pl2:PORNAH1 | + | 29.00 | 130.27 | 342 | Y14556 Pleurotus ostreatus mRNA |
| gb_sts:G46114 | + | 29.00 | 127.13 | 485 | G46114 26550_1 zebrafish AB Danio |
| gb_pl1:POA5060 | + | 29.00 | 126.54 | 518 | AJ225060 Pleurotus ostreatus PC |
| gb_ba2:AF109102 | - | 29.00 | 125.45 | 585 | AF109102 Haemophilus influenzae |
| gb_ba1:LPY15954 | - | 29.00 | 115.37 | 1797 | Y15954 Lactobacillus plantarum |
| gb_vl:ALV1828 | + | 29.00 | 115.21 | 214.66 | X87254 Artichoke italian latex |
| gb_in1:AF020343 | + | 29.00 | 112.17 | 317.11 | AF020343 Caenorhabditis elegans |
| gb_in1:AF020342 | + | 29.00 | 112.12 | 319.10 | AF020342 Caenorhabditis elegans |
| gb_ba1:NE2357115 | - | 29.00 | 110.90 | 373.26 | X87286 N. exedens 23S rRNA gene |
| gb_in2:AF032112 | - | 29.00 | 110.66 | 384.92 | AF032112 Caenorhabditis elegans |
| gb_vl:HSIRLB123 | + | 29.00 | 110.31 | 402.80 | X82333 H. sapiens IRLB gene (ex |
| gb_vl:PXVMEV1A | + | 29.00 | 110.25 | 405.88 | M85588 Pigeon pox virus major |
| gb_pl1:YS2BGLUA | + | 29.00 | 110.14 | 411.33 | M2475 S. fibuligera extracellu |
| gb_pl1:AB0189511 | + | 29.00 | 110.11 | 413.10 | AB018951 Homo sapiens PACE4 ge |
| gb_pl1:YS2BGLUB | + | 29.00 | 109.72 | 434.40 | M2476 S. fibuligera extracellu |
| gb_vl:FPVMEV1A | + | 29.00 | 109.14 | 467.82 | M85587 Fowlpox virus major env |
| gb_pl1:SPC0C7 | - | 29.00 | 106.86 | 626.57 | X78799 S. pombe (wildtype 972 H |
| gb_ba1:AB024550 | + | 29.00 | 106.51 | 655.31 | AB024550 Bacillus halodurans q |
| gb_ba1:PWRCIT | + | 29.00 | 106.48 | 657.67 | M11559 Plasmid pWR60 (from E.c |
| gb_ba1:ECOPUV | + | 29.00 | 106.27 | 675.88 | D90281 Escherichia coli pepP, |
| gb_ba1:BSGENDNA | + | 29.00 | 105.75 | 722.82 | Z34287 B. subtilis (Soil3) geno |
| gb_ba2:AE000341 | - | 29.00 | 99.75 | 1.6e+03 | AE000341 Escherichia coli K-1 |
| gb_ba2:AE000374 | - | 29.00 | 99.40 | 1.6e+03 | AE000374 Escherichia coli K-1 |
| gb_ba2:AE000312 | - | 29.00 | 98.15 | 1.9e+03 | AE000312 Escherichia coli K-1 |
| gb_ba2:AE000955 | + | 29.00 | 96.46 | 2.4e+03 | AE000955 Archaeoglobus fulgid |
| gb_pl1:SPB021 | + | 29.00 | 96.29 | 2.4e+03 | AL035537 S. pombe chromosome I |
| gb_ba1:D90854 | - | 29.00 | 95.12 | 2.8e+03 | D90854 E. coli genomic DNA, KC |
| gb_pl1:SPB017186 | + | 29.00 | 94.80 | 2.9e+03 | AC017186 Drosophila melanogas |
| gb_pl1:SPB012D12 | + | 29.00 | 94.67 | 3.0e+03 | AL035085 S. pombe chromosome I |
| gb_ba1:D90854 | + | 29.00 | 94.63 | 3.0e+03 | D90854 E. coli genomic DNA, KC |
| gb_in1:CEL81348 | + | 29.00 | 94.18 | 3.2e+03 | AF039717 Caenorhabditis eleg |
| gb_ba1:SPB03XSE | + | 29.00 | 90.71 | 5.0e+03 | Z70177 B. subtilis DNA (28 kb |
| gb_ba2:AC014909 | + | 29.00 | 89.70 | 5.7e+03 | AC014909 Drosophila melanogas |
| gb_pl1:D89861 | - | 29.00 | 89.45 | 5.8e+03 | D89861 Cyanidochydon merola |
| gb_in1:CEK02E11 | - | 29.00 | 89.24 | 6.0e+03 | Z77665 Caenorhabditis elegans |
| gb_in1:CEK08C7 | - | 29.00 | 88.93 | 6.3e+03 | Z70286 Caenorhabditis elegans |
| gb_ba2:AC009604 | + | 29.00 | 88.45 | 6.6e+03 | AC009604 Leishmania major chr |
| gb_in1:CEC12D8 | + | 29.00 | 88.44 | 6.7e+03 | Z73969 Caenorhabditis elegans |
| gb_in1:CEK10E8 | - | 29.00 | 87.85 | 7.2e+03 | Z81576 Caenorhabditis elegans |
| gb_ba2:HSU71217 | + | 29.00 | 87.52 | 7.5e+03 | U71217 Human clone c15H12, 24 |
| gb_ba2:AF102543 | + | 29.00 | 87.25 | 7.7e+03 | AF102543 Zymomonas mobilis 2M |
| gb_pr3:HSN80H12 | + | 29.00 | 86.84 | 8.2e+03 | Z80902 Human DNA sequence fro |

gb_htg7:AC018207 - 29.00 86.47 8.6e+03 44887 ! AC018207 Drosophila melano
gb_htg3:AC008807 + 29.00 84.26 1.1e+04 57376 ! AC008807 Homo sapiens chro
gb_htg3:AC008681 - 29.00 84.10 1.2e+04 58456 ! AC008681 Homo sapiens chro

seq_name: gb_pl2:PORNAH1

seq_documentation_block: 342 bp mRNA PLN 17-AUG-1999

LOCUS PORNAH1

DEFINITION Pleurotus ostreatus mRNA for hydrophobin 1.

ACCESSION Y14556

VERSION Y14556.1 GI:2370368

KEYWORDS hydrophobin; POH1 gene.

SOURCE oyster mushroom.

ORGANISM Pleurotus ostreatus

Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Aphyllophorales;

Lentinaceae; Pleurotus.

REFERENCE 2 (bases 1 to 342)

AUTHORS Asgeirsdottir, S.A., de Vries, O.M. and Wessels, J.G.

TITLE Identification of three differentially expressed hydrophobins in

Pleurotus ostreatus (oyster mushroom)

Microbiology 144 (Pt 11), 2961-2969 (1998)

MEDLINE 99061188

REFERENCE 2 (bases 1 to 342)

AUTHORS Asgeirsdottir, S.A.

TITLE Direct Submission

JOURNAL Submitted (22-AUG-1997) S.A. Asgeirsdottir, Department of Plant

Biology, University of Groningen, Kercklaan 30, 9751 NN Haren,

NETHERLANDS

FEATURES

Location/Qualifiers

1..342

/organism="Pleurotus ostreatus"

/db_xref="taxon:5322"

/dev_stage="fruiting"

/tissue_type="fruit body"

sig_peptide 1..75

/gene="POH1"

1..342

/gene="POH1"

/codon_start=1

/product="hydrophobin"

/protein_id="CAA74986.1"

/db_xref="GI:2370368"

/db_xref="SPTREMBL:O13502"

/translation="MESIRISTVLAASALLAVAIPTMTETPCNTGPIQCCNSVOS

ATSSAAGAPLAAGLVLSGLASLVGLGSLQVIGVGANSQSSQAACCTGNTFNGA

VVLGCSPIKLL"

1..342

/gene="POH1"

76..339

/gene="POH1"

BASE COUNT 61 a 114 c 83 g 84 t

ORIGIN

alignment_scores:

Quality: 29.00

Ratio: 4.833

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 6

caps: 0

alignment_block:

US-08-653-294-7 x PORNAH1 ..

Align seg 1/1 to: PORNAH1 from: 1 to: 342

1 TyrArgLeuAlaileArg 6

|||||

186 TATCGCGCTGCTATTAGG 203

seq_name: gb_sts:G46114

seq_documentation_block: 485 bp DNA

LOCUS G46114

DEFINITION Z6550_1 zebrafish AB Danio rerio STS genomic clone Z6550 5',

23-MAR-1999


```

exon          /number=2
259..420
/feature="POH1"
intron        /number=3
421..473
/feature="POH1"
exon          /number=3
474..518
/feature="POH1"
/number=4
97 a 168 c 121 g 132 t

BASE COUNT
ORIGIN

alignment_scores:
  Quality: 29.00      Length: 6
  Ratio: 4.833       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-7 x POAJ5060 ..
Align seg 1/1 to reverse of: AF109102 from: 1 to: 518

seq_name: gb_ba2:AF109102

seq_documentation_block:
LOCUS AF109102 585 bp DNA BCT 08-APR-1999
DEFINITION Haemophilus influenzae UC18 outer membrane protein 26 (skp) gene,
partial cds.
ACCESSION AF109102
VERSION AF109102.1 GI:4574279
KEYWORDS
SOURCE Haemophilus influenzae.
ORGANISM Haemophilus influenzae.
REFERENCE 1 (bases 1 to 585)
AUTHORS El-Adhami,W., Kyd,J.M., Bastin,D.A. and Cripps,A.W.
TITLE Characterization of the gene encoding a 26-kilodalton protein
(OMP26) from nontypeable Haemophilus influenzae and immune
responses to the recombinant protein
JOURNAL Infect. Immun. 67 (4), 1935-1942 (1999)
MEDLINE 99185023
REFERENCE 2 (bases 1 to 585)
AUTHORS El-Adhami,W., Kyd,J.M., Bastin,D.A. and Cripps,A.W.
TITLE Direct Submission
JOURNAL Submitted (25-NOV-1998) Human and Biomedical Sciences, University
of Canberra, Canberra, ACT 2601, Australia
FEATURES
  source
    1..585
    /organism="Haemophilus influenzae"
    /strain="UC18"
    /db_xref="taxon:727"
    <1..582
    /gene="skp"
    <1..582
    /gene="skp"
    /note="OMP26: Skp"
    /codon_start=1
    /transl_table=11
    /product="outer membrane protein 26"
    /protein_id="AAD23984.1"
    /db_xref="GI:4574280"
    /translation="ALALGALASGYAAAEKIAFAINAGYIFQNHDPDQRAVDKLDAA
FKPVAEKLAASKVEVDKIAAARKKVEAKVALEKAPRLQRQADIKRQEEINKLGA
EDAELOKMQEOKVEQAEERKQLDSIQATNNLAKARKAGTYTVLDAN
SVYFAVEGDKITEVLKSIPEAKAAAEKK"

BASE COUNT 229 a 93 c 126 g 137 t

ORIGIN

alignment_scores:
  Quality: 29.00      Length: 6
  Ratio: 4.833       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-7 x POAJ5060 ..
Align seg 1/1 to reverse of: AF109102 from: 1 to: 518

seq_name: gb_ba2:AF109102

seq_documentation_block:
LOCUS AF109102 585 bp DNA BCT 08-APR-1999
DEFINITION Haemophilus influenzae UC18 outer membrane protein 26 (skp) gene,
partial cds.
ACCESSION AF109102
VERSION AF109102.1 GI:4574279
KEYWORDS
SOURCE Haemophilus influenzae.
ORGANISM Haemophilus influenzae.
REFERENCE 1 (bases 1 to 585)
AUTHORS El-Adhami,W., Kyd,J.M., Bastin,D.A. and Cripps,A.W.
TITLE Characterization of the gene encoding a 26-kilodalton protein
(OMP26) from nontypeable Haemophilus influenzae and immune
responses to the recombinant protein
JOURNAL Infect. Immun. 67 (4), 1935-1942 (1999)
MEDLINE 99185023
REFERENCE 2 (bases 1 to 585)
AUTHORS El-Adhami,W., Kyd,J.M., Bastin,D.A. and Cripps,A.W.
TITLE Direct Submission
JOURNAL Submitted (25-NOV-1998) Human and Biomedical Sciences, University
of Canberra, Canberra, ACT 2601, Australia
FEATURES
  source
    1..585
    /organism="Haemophilus influenzae"
    /strain="UC18"
    /db_xref="taxon:727"
    <1..582
    /gene="skp"
    <1..582
    /gene="skp"
    /note="OMP26: Skp"
    /codon_start=1
    /transl_table=11
    /product="outer membrane protein 26"
    /protein_id="AAD23984.1"
    /db_xref="GI:4574280"
    /translation="ALALGALASGYAAAEKIAFAINAGYIFQNHDPDQRAVDKLDAA
FKPVAEKLAASKVEVDKIAAARKKVEAKVALEKAPRLQRQADIKRQEEINKLGA
EDAELOKMQEOKVEQAEERKQLDSIQATNNLAKARKAGTYTVLDAN
SVYFAVEGDKITEVLKSIPEAKAAAEKK"

BASE COUNT 229 a 93 c 126 g 137 t

ORIGIN

alignment_scores:
  Quality: 29.00      Length: 6
  Ratio: 4.833       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-7 x AFI09102/rev ..
Align seg 1/1 to reverse of: AFI09102 from: 1 to: 585

seq_name: gb_bal:LPX15954

seq_documentation_block:
LOCUS LPX15954 1797 bp DNA BCT 02-JUL-1998
DEFINITION Lactobacillus plantarum bglT gene.
ACCESSION Y15954
VERSION Y15954.1 GI:3288505
KEYWORDS beta-glucosidase; bglT gene.
SOURCE Lactobacillus plantarum.
ORGANISM Lactobacillus plantarum
Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
Lactobacillus.
REFERENCE 1 (bases 1 to 1797)
AUTHORS Marasco,R., Muscarello,L., Varcamonti,M., De Felice,M. and
Sacco,M.
TITLE Expression of the bglH gene of Lactobacillus plantarum is
controlled by carbon catabolite repression
JOURNAL J. Bacteriol. 180 (13), 3400-3404 (1998)
MEDLINE 98317280
REFERENCE 2 (bases 1 to 1797)
AUTHORS Sacco,M.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-1997) M. Sacco, Istituto Internazionale di
Genetica e Biofisica, via Marconi 10, I- 80125, Napoli, ITALY
FEATURES
  source
    1..1797
    /organism="Lactobacillus plantarum"
    /strain="B21"
    /db_xref="taxon:1590"
    157..162
    -35_signal
    181..186
    -10_signal
    190..203
    enhancer
    235..1677
    gene
    235..1677
    /gene="bglT"
    /note="cyclic-AMP responsive element (CRE)"
    /codon_start=1
    /transl_table=11
    /product="beta-glucosidase"
    /protein_id="CA375906.1"
    /db_xref="GI:3288506"
    /db_xref="SPTREMBL:O86291"
    /translation="MSEFFPEGLWGATAANQLEGQYQEGGRLSIADRLPGGKDRFK
IVSQPDFWTIDTKYTPNHEGIDFYHHYKEDIALFAEMGKCYRFSIAWSRIFFNG
DETQNEAGLKFYDDVIDECLANNIEPVITISHVELPLNLAKRYGKNRYLIEFVET
FARTILTRASKYKVMTEINSAYHFFVPMGGLVLSLGANDKKNVQFNWHQVFAS
AKAVIAHELRDDIQVGMILYATSDNSPNVQLANLOHNODFNFFCADQVQRGAY
PYTKRLAEYNLTFDDETDGLALQYPVDYIGFSYMSAVETGTSVTDVAG
NLMGKYNPFLKASDWGIDPTGLRIALNELHDYKQKPLFVVENGIGAIKDKPKNFY
VEDDYRIDYVKQHIEMAGAIIDDGVLMGYTPWGCIDILVSASTGSEMSKRYGFIYVDLD
DQHGTLARYPKKSFYQDVIKHSLTKK"
    1696..1732
    terminator
    1696..1732
    BASE COUNT 520 a 305 c 438 g 534 t
    ORIGIN
```

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-08-653-294-7 x LPV15954/rev ..

Align seg 1/1 to reverse of: LPV15954 from: 1 to: 1797

1 TyrArgLeuAlaIleArg 6
 |||||

928 TATCGTCTCGCAATCGT 911

seq_name: gb_vi:A1LV1828

seq_documentation_block:
 LOCUS A1LV1828 1828 bp mRNA VRL 19-FEB-1996
 DEFINITION Artichoke Italian latent virus mRNA unknown function (1828bp).
 ACCESSION X87254
 VERSION X87254.1 GI:1199791
 KEYWORDS
 SOURCE
 ORGANISM
 artichoke Italian latent virus.
 artichoke Italian latent virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;
 Nepovirus.

REFERENCE 1 (bases 1 to 1828)

AUTHORS Grieco.F.

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1828)

AUTHORS Grieco.F.

TITLE Direct Submission

JOURNAL Submitted (05-MAY-1995) F. Grieco, Univ. of Bari, Dipt. Protezione

Plante, via Amendola 165/A, 70126 Bari, ITALY

FEATURES

Location/Qualifiers

1..1828

/organism="artichoke Italian latent virus"

/db_xref="taxon:46075"

1..1534

/gene="ORF"

<1..1534

/gene="ORF"

/function="unknown"

/codon_start=2

/protein_id="CAA60707.1"

/db_xref="GI:1199792"

/db_xref="SPTREMBL:Q64959"

/translation="NGDFAFSQRITVPAARTVGTIGTIDIPALITTTNSRVCAEWLE

RGYVDRNLWVSHLTPSYLGMALWVFDAGHPTDVTITVELESIRHLSPHVHLK

DNVTSTWLLNFHREGQSILNFAFGPMKPKVWIIAASQAQPCSDADVOYVVEGYPOEK

VSEDLRLKRKILITLVESTHLDLLOLLAPQOLAIGTATNPFPSAEKSTSTKRE

TTYSAGLLSHFLGIGGLKRFVHSTSSCLLTSKURVFNHESQPCRRPAPHTUTLM

GLDWRAVQAACATNFAGSGAREFWIMPISAPRAPQVETKFEYIRILGDIVPDL

CROINYKORFWMISPSDKTTTDLDFKIPSRIGNIGVKNKCVNFTNAFALMCATIG

MHWGRCILHFTWSHRNTEAGKNAREISQSRLEWVILQHLILGIRGLIVVDNAYSIP

FEFGSFAGPVISGGTPNEARNVRVQSTSWQIHAVTVSIEVLPGRFYGRSAGPMTI

PS"

alignment_scores:

Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-7 x A1LV1828 ..

Align seg 1/1 to: A1LV1828 from: 1 to: 1828

1 TyrArgLeuAlaIleArg 6

|||||

1047 TATCGCTACGATAGA 1064

seq_name: gb_inl:AF020343

seq_documentation_block:

LOCUS AF020343 2565 bp mRNA INV 15-NOV-1997
 DEFINITION Caenorhabditis elegans fork head-related transcription factor

DAF-16a2 (daf-16) mRNA, complete cds.

ACCESSION AF020343

VERSION AF020343.1 GI:2618978

KEYWORDS

SOURCE

ORGANISM

Caenorhabditis elegans.

Caenorhabditis elegans.

Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

1 (bases 1 to 2565)

AUTHORS Ogg.S., Paradis.S., Gottlieb.S., Patterson.G.I., Lee.L.,

Tissenbaum.H.A. and Ruvkun.G.

TITLE The Fork head transcription factor DAF-16 transduces insulin-like

metabolic and longevity signals in C. elegans

Nature 389 (6654), 994-999 (1997)

JOURNAL 98013175

REFERENCE 2 (bases 1 to 2565)

AUTHORS Ogg.S., Paradis.S. and Ruvkun.G.

TITLE Direct Submission

JOURNAL Submitted (21-AUG-1997) Molecular Biology, MGH, 50 Blossom St.,

Boston, MA 02114, USA

FEATURES

Location/Qualifiers

1..2565

/organism="Caenorhabditis elegans"

/strain="N2"

/db_xref="taxon:6239"

/map="1; between mgp45 and mgp49"

1..2565

/gene="daf-16"

/note="null phenotype is suppression of daf-2"

297..1823

/gene="daf-16"

/codon_start=1

/product="fork head-related transcription factor DAF-16a2"

/protein_id="AAB84391.1"

/db_xref="GI:2618979"

/translation="MMEMLVQGTDASSASTSSVSRECGATFMNTPDDVMNDMM

EPFPRCRNTWPMRPPQLPEPNSPIHQEIPEDADLIGNSCCGCGGASNGST

AMLHTPDGNSHQTSPFSESESPDDTVSGKKTTRRNAMNMSYAEILITAIMASPE

KRLTLAQVYEWVQVYFRDKGDSNSAGKNSIRHNSLSHFRMIRIQEGAGKSSW

WVINPDARQPNRPTRETSIETTTKAQLKSRGAKKRIKRALMGSLHTLNGN

SIAGSIOTISHDLDDDSMOGAFNVSPSPRTQSNLSIPGSSSRVSPAGSDIYDD

LEFSSWGESVPAIPSDIVDTDMRIDATTHIGVQIKQESKPIKTEPIAPPSYHE

LNSVSGCAQNPILLRNIVPSTNPKPPLGAYGNQNGGTPINWLSTSSNSPLPGI

QSCVGRQAQHTVASSSALLIDLEMLTLPDQPLMDTMDVALIRHLSQAGQGHIFDL

alignment_scores:

Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-7 x AF020343/rev ..

Align seg 1/1 to reverse of: AF020343 from: 1 to: 2565

1 TyrArgLeuAlaIleArg 6

|||||

681 TATCGTCTCGCATCGG 664

seq_name: gb_inl:AF020342

```

seq_documentation_block: 2579 bp mRNA INV 13-NOV-1997
LOCUS AF020342 Caenorhabditis elegans fork head-related transcription factor
DEFINITION DAF-16a1 (daf-16) mRNA, complete cds.
ACCESSION AF020342
VERSION AF020342.1 GI:2618976
KEYWORDS
SOURCE
ORGANISM
Caenorhabditis elegans.
Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 2579)
Ogg, S., Paradis, S., Gottlieb, S., Patterson, G.I., Lee, L.,
Tissenbaum, H.A. and Ruvkun, G.
The Fork head transcription factor DAF-16 transduces insulin-like
metabolic and longevity signals in C. elegans
Nature 389 (6654), 994-999 (1997)
98013175
REFERENCE
2 (bases 1 to 2579)
Ogg, S., Paradis, S. and Ruvkun, G.
Direct Submission
TITLE
The Fork head transcription factor DAF-16a1
JOURNAL
Submitted (21-AUG-1997) Molecular Biology, MGH, 50 Blossom St.,
Boston, MA 02114, USA
FEATURES
source
1..2579
Location/Qualifiers
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/chromosome="1"
/map="1; between mgp45 and mgp49"
1..2579
/gene="daf-16"
/note="null phenotype is suppression of daf-2"
299..1831
/gene="daf-16"
/codon_start=1
/product="fork head-related transcription factor DAF-16a1"
/protein_id="AAB84390.1"
/db_xref="GI:2618977"
/translation="MEMLVQGTASSSTSTSVSRFGADTFMNTDPDVMNDMM
EPIPRDRCNTWPMRPOLEPLNSSPIIHEQIPEEDADLYSNEQQLGASSNGST
AMLHTPDGNSHSTSPDFRMSPEPDVTYSGKTTTNRNAGNMSYAEIITAIMAS
PERKLLAQVYEMWQVNPYFRDKGDSNSAGWKNISIRHNLSLHSRMRIONEGAGKS
SWVINDPAKPGNPRTRERSNTIETTKAQLKSRGAKKRIKERALMGSLHSTLIN
GNSIAGSIQIISHDLDDSMQAFDNVPSFRQSNLSIPGSSSRVSPAIGSDIY
DDEFFSWGESVPAIPSDIVDTQMRIDATTHIGVQIKESKPIKTEPIAPPSY
HELNSVSGCAQPLLRNPVPTNFKMPLPGAYNGYQNGGTPINWLSNSSLPL
GIQSCGIVAAQHTVASSSALPDLNLTLPDQPLMDTMDVDAIRHELSSQAGQHIF
DL"
BASE COUNT 686 a 645 c 461 g 787 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-7 x NE23S7ITS/rev ..
Align seg 1/1 to reverse of: NE23S7ITS from: 1 to: 2955

1 TyrArgLeuAlaIleArg 6
|||||
2923 TACCGGTAGCTATACGC 2906

seq_name: gb_in2:AF032112
seq_documentation_block:
LOCUS AF032112 3035 bp mRNA INV 19-NOV-1997
DEFINITION Caenorhabditis elegans DAF-16 (daf-16) mRNA, complete cds.
ACCESSION AF032112
VERSION AF032112.1 GI:2623942
KEYWORDS
SOURCE
Caenorhabditis elegans.
Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 3035)
Lin, K., Dorman, J.B., Rodan, A. and Kenyon, C.
daf-16: An HNF-3/forkhead family member that can function to double
the life-span of Caenorhabditis elegans
Science 278 (5341), 1319-1322 (1997)
98028757
REFERENCE
2 (bases 1 to 3035)

```

AUTHORS Lin, K., Dorman, J.B., Rodan, A. and Kenyon, C.
 TITLE Direct Submission
 JOURNAL Submitted (30-OCT-1997) Biochemistry and Biophysics, UCSF, 513
 Panassus Ave., San Francisco, CA 94143-0554, USA

FEATURES
 source
 1. .3035
 /organism="Caenorhabditis elegans"
 /strain="Bristol N2"
 /db_xref="taxon:6239"
 /chromosome="I"
 gene
 1. .3035
 /gene="daf-16"
 /note="defective dauer formation; suppresses Age phenotype
 of age-1, daf-2."
 334. .1866
 /gene="daf-16"
 /note="HNF-3/forkhead homolog"
 /codon_start=1
 /product="DAF-16"
 /protein_id="AAC47803.1"
 /db_xref="GI:2623943"
 /translation="MMELVDOGTDASSASTSTSVSRFGADTFMTNPDVMMDDM
 EPIPRDCTWPMRRPQLEPPLNSPIIHEQIPEEDADLYGSGQGLGASSNGST
 AMLPDQSGNSHOTFPFSDRMSQPDVTGSKTTTRNAGNMSYAEIITAIWAS
 PKRLTIAOVYEMVQNVYFPRDKGDSNSGKNSIRKNSLHSHRMRIONEGAGKS
 SMWVNPDAKPGENRPRPERSNTIETTKAOLKSRKAKKRIKRALMSLHSTLN
 GNSIAGSIOTISHDLYDDSMQAGDNFSPSRPTQSNLSIPGSSRSVSPAIGSDIY
 DLEFPVSGESVPAIPSDIVDTQMDRATTHGGVQIQKSPKTEPIAPPPSY
 HELNSVSGCAQPLLRNPVPTNFKPMPLFGAYNGYGGITPINLSTNSPPLP
 GIQSGGIVAAQHTVASSALPDLLENLTLPDQMDTMDVDALIRHLSQAGGQHIHF
 DL"
 BASE COUNT 796 a 759 c 511 g 969 t
 ORIGIN

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-08-653-294-7 x AF032112/rev ..
 Align seg 1/1 to reverse of: AF032112 from: 1 to: 3035

1 TyrArgLeuAlaIleArg 6
 |||||
 724 TATCGTCTGGCGATTCGG 707

seq_name: gb_pri:HSIRLB123

seq_documentation_block:
 LOCUS HSIRLB123 3157 bp DNA PRI 18-NOV-1994
 DEFINITION H.sapiens IRLB gene (exon1-3).
 ACCESSION X82333
 VERSION X82333.1 GI:562754
 KEYWORDS DNA-binding protein; ir1b gene.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 3157)
 AUTHORS Stasiv, Y.Z. and Itkes, A.V.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3157)
 AUTHORS Stasiv, Y.Z.
 TITLE Direct Submission
 JOURNAL Submitted (26-OCT-1994) Y.Z. Stasiv, Engelhardt Institute of
 Molecular Biol., Vavilov Str. 32, 117984, Moscow B-334, RUSSIA
 COMMENT Related sequence: X63417.
 FEATURES
 Location/Qualifiers
 1. .3157
 /organism="Homo sapiens"

/db_xref="taxon:9606"
 /cell_line="Hela"
 1. .166
 /number=1
 /evidence=experimental
 join(24. .166,2151. .2268,2891. .3054)
 /gene="ir1b"
 /function="putative transcription factor"
 /product="DNA-binding protein"
 24. .3054
 /gene="ir1b"
 167. .2150
 /gene="ir1b"
 /number=1
 /evidence=experimental
 2151. .2268
 /gene="ir1b"
 /number=2
 /evidence=experimental
 2269. .2890
 /gene="ir1b"
 /number=2
 /evidence=experimental
 2891. .3053
 /gene="ir1b"
 /number=3
 /evidence=experimental
 3054. .>3157
 /number=3
 /evidence=experimental
 BASE COUNT 804 a 579 c 583 g 1183 t 8 others
 ORIGIN

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-08-653-294-7 x HSIRLB123 ..
 Align seg 1/1 to: HSIRLB123 from: 1 to: 3157

1 TyrArgLeuAlaIleArg 6
 |||||
 647 TACAGGCTAGCAATAGA 664

seq_name: gb_v1:PXVMEVNA

seq_documentation_block:
 LOCUS PXVMEVNA 3178 bp DNA VRL 03-AUG-1993
 DEFINITION Pigeon pox virus major envelope antigen (p37K) gene, 3' end cds;
 envelope protein (p74K) gene, complete cds; ORF, 5' end cds.
 ACCESSION M88588
 VERSION M88588.1 GI:333522
 KEYWORDS envelope protein; homologue; major envelope antigen.
 SOURCE Pigeon pox virus DNA.
 ORGANISM Pigeon pox virus
 Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 Avipoxvirus.
 REFERENCE 1 (sites)
 AUTHORS Goebel, S.J., Johnson, G.P., Perkins, M.E., Davis, S.W., Winslow, J.P.
 and Paoletti, E.
 TITLE The complete DNA sequence of Vaccinia virus
 JOURNAL Virology 179, 247-266 (1990)
 MEDLINE 91021027
 REFERENCE 2 (sites)
 AUTHORS Calvert, J.G., Ogawa, R., Yanagida, N. and Nazerian, K.
 TITLE Identification and functional analysis of the fowlpox virus homolog
 of the vaccinia virus p37K major envelope antigen gene
 JOURNAL Virology 191, 783-792 (1992)
 MEDLINE 93079881

```
REFERENCE 3 (bases 1 to 3178)
AUTHORS Ogawa,R., Yanagida,N., Nazerian,K. and Calvert,J.G.
TITLE Insertional inactivation of a fowpox virus homologue of the
vaccinia virus F12L gene inhibits release of enveloped virions
J. Gen. Virol. 74, 55-64 (1993)
MEDLINE 93139784
FEATURES
    source
        Location/Qualifiers
            1..3178
                /organism="Pigeonpox virus"
                /db_xref="taxon:10264"
            1..756
                /gene="p43K"
            <1..756
                /gene="p43K"
            /note="homologue to vaccinia virus p37K (F13L) gene;
            putative"
            /citation=[1]
            /codon_start=1
            /product="major envelope antigen"
            /protein_id="AAA47186.1"
            /db_xref="GI:333523"
            /translation="KLHFEYIGSASLMGNALTIKNGIYSENNSLAMDLYFRSLDYKI
            ISKKCLFTRMATKYHFKFNKNGIFFSDSPHVMGKRTFGLDCVHIHVIDAKSTID
            LAISLLPKTKTSIVYPIIKDALIRAVLGRVGLRVLGLFWKTKDVIDKASIKSL
            NELGVHDIDISTKVFPPVNSKVDINNSKMMIIDGRYAHVNTANLDGSHFNHFAFVS
            FNCMDQOFTKKAIEVFERDWSIPYAKEIDMSQI"
            794..2686
                /gene="p74K"
            794..2686
                /gene="p74K"
            /note="disruption of this gene results in a small plaque
            phenotype; homolog of vaccinia virus F12L; putative"
            /citation=[1]
            /codon_start=1
            /function="Involved in the release of progeny virions"
            /protein_id="AAA47187.1"
            /db_xref="GI:333524"
            /translation="NALIEQLQSSEQSILSPRYKFKDFHNVIFTTIDETLIVITY
            NVPLRLAITEKITEFPFSTGCIITSNNSDIDTDYFIPNSLSLLDLKKRAYD
            NLRDLFSAMSEMNDELNSDIYSLNKLHLKHLNLLYKLVLSIDIDRYKLNKKN
            TIDVISVNGRNINWIKVDIYIETIYSEYLRWSDIKRATESNWLPIQSINPLNENI
            YAFETALERSNERLNTGAIELYPDIIITGRNNEIDIEKFLDQLEEVYIKNSDSIV
            LTGYHLTLENTILERYISKYDWITCNRLVCKTGTVEVLFDAAIFFPSSNKKGY
            KWTGKLNKFNKFKFQSOLEKYINNNSVAERIYYQLSLRKHHSICLIEIFELNGDF
            NPSGLDLILFISIRVKNNGNYYYPKSSAVNMLSSIYTDVYAIIDDIDKSKKLVEN
            SFPLIMSGYPEGKPYTKPEKGYLSICLDVEISNDIKNPILYCKENKSKRFTG
            VFTSDVIDTAVKRGYKKILECIENPNKIKLFEDNICYLNKLFIEHODYTHDEKSLQ
            YLFSYLLKGNVTEDVLAMSKRNLSIISFIISYCRNNTYIKLECPVYESSNIVKCKY
            NQVIYK"
            2726..>3178
                /note="ORF"
            /codon_start=1
            /function="unknown"
            /protein_id="AAA47188.1"
            /db_xref="GI:333525"
            /translation="MDTNKRSLDEHDTGEESPGKLIQIVEINDEEDINFTDNPYKLV
            KSRDINSILVPLVGVMIKINDIKGVTDKVNLKLPKTSKTSNSTSCINPIDSIPNF
            LDDGKNFYNSVEVSILOVSHGNDMNIDKYDGSFDYAVLCLEKSGRS"
BASE COUNT 1232 a 406 c 492 g 1048 t
ORIGIN

alignment_scores:
    Quality: 29.00 Length: 6
    Ratio: 4.833 Gaps: 0
    Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
    US-08-653-294-7 x PXVMENVA/rev ..
    Align seg 1/1 to reverse of: PXVMENVA from: 1 to: 3178
    1 TyrArgLeuAlaIleArg 6
    |||||
    513 TACCGGTGGCAATTAGA 530
seq_name: gb_prl:AB00189511
```

```
|||||
1426 TATAGACTGGCTATACGG 1409
seq_name: gb_prl:YS2BGLUA

seq_documentation_block:
    LOCUS YS2BGLUA 3215 bp DNA PLN 27-APR-1993
    DEFINITION S.fibuligera extracellular beta-glucosidase 1 (BGL1) gene, complete
    cds
    ACCESSION M22475
    VERSION M22475.1 GI:170807
    KEYWORDS BGL1 protein; beta-glucosidase 1.
    SOURCE S.fibuligera (strain HUT77212) DNA, clones pSFCB[1,2,3,4,5].
    ORGANISM Saccharomycopsis fibuligera
    Eukaryota; Fungi; Ascomycota; Saccharomycetales;
    Saccharomycopsidaceae; Saccharomycopsis.
    REFERENCE 1 (bases 1 to 3215)
    AUTHORS Machida,M., Ohtsuki,I., Fukui,S. and Yamashita,I.
    TITLE Nucleotide sequences of Saccharomycopsis fibuligera genes for
    extracellular beta-glucosidases as expressed in Saccharomyces
    cerevisiae
    JOURNAL Appl. Environ. Microbiol. 54, 3147-3155 (1988)
    MEDLINE 89133518
    COMMENT Draft entry and sequence for [1] kindly submitted by I.Yamashita,
    09-FEB-1989.
    FEATURES
        Location/Qualifiers
            1..3215
                /organism="Saccharomycopsis fibuligera"
                /db_xref="taxon:4944"
            277..327
                /note="beta-glucosidase 1 signal peptide"
            277..2907
                /note="beta-glucosidase 1 precursor"
            /codon_start=1
            /protein_id="AAA34314.1"
            /db_xref="GI:170808"
            /translation="MLMIVQLLVFALGLAVAVPIQNTQSPQRDESSQVSPHYTPT
            POGGLQDWQEAARAKAIVGOMTIVEKVNLTGTGQWLDPCVGTGTSVPRGIPNL
            CLOGPLAVRADFTVGYPSGLATGATNKDLFLQGOALGHEFNKGVHIALGPAVG
            PLGVKARGRFEAFGSDPYLQGTAAATIKGLQNNYMACVHKHFIHQEQYRQDD
            INPATYTKAISAINRQWHAHALYLPFADSVRAGVSVKCSNVRVNNNTACENSY
            MNHLLEKPELQFQVSDWGAQSGVYSAISGLDMSMPGEVYGGWNTGTGFWGNLT
            KALNYETPIERLDDMTAIIAALYATNSFTDLHPNFSWTTKEYGNKYADNTTE
            IVKYNVDPNSDFTDLKVAEESIVLLKNENLTPI SPEKAKRLLSLGGTAAGDPD
            IGYCEDOSCTNGALFOGWSGSGVQYVTPPEEISYLARKNKWQDYIRESYDLA
            QVTVASDAHLISIVVVSASSEGIVTDCNODKRNILNNGDKLIETVABNCANTV
            VVVSTGQINFEFGADHFNVTAVWAGPLGDSRGTAINILFGKRNPSCHLPFTTAKT
            DDDYIPITYSPSSGEPEDNHLVDLLDYRYFEENIEPRYAFYGLSYGLNEYEVSN
            AKVSAKVDDEELPEPATYLSFESYQNAKDSKPSDAFAPADLRNRYLILPYLDSNY
            TLKGNIEYDPGYSTQRTTPNQGGLGNDALWEVAYNSTKDFVPGNSTDKFVPO
            LYLKHPDGKFTPIQLRGFEKVELSPGEKKTVDLRLRLDLSVNDTTRQSVIBSGT
            YEALIGVAVNDIKTSVLEFI"
            328..2904
                /note="beta-glucosidase 1"
BASE COUNT 969 a 614 c 667 g 965 t
ORIGIN 1 bp upstream of BsteII site.

alignment_scores:
    Quality: 29.00 Length: 6
    Ratio: 4.833 Gaps: 0
    Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
    US-08-653-294-7 x YS2BGLUA ..
    Align seg 1/1 to: YS2BGLUA from: 1 to: 3215
    1 TyrArgLeuAlaIleArg 6
    |||||
    513 TACCGGTGGCAATTAGA 530
seq_name: gb_prl:AB00189511
```

```

seq_documentation_block:
LOCUS AB00189511 3227 bp DNA PRI 13-FEB-1999
DEFINITION Homo sapiens PACE4 gene, exon 14, 15.
ACCESSION AB001908
VERSION AB001908.1 GI:2281763
KEYWORDS PACE4; alternative splicing.
SEGMENT 11 of 17
SOURCE Homo sapiens DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3227)
AUTHORS Matsuda,Y.
TITLE Direct Submission
JOURNAL Submitted (14-MAR-1997) to the DDBJ/EMBL/GenBank databases. Yoshiko
Matsuda, The University of Tokushima, Department of Biological
Science and Technology; Minamijosanjima-cho 2-1, Tokushima,
Tokushima 770, Japan [E-mail:matsuda@bio.tokushima-u.ac.jp,
Tel:0886-56-7523, Fax:0886-55-3161]
REFERENCE 2 (sites)
AUTHORS Tsuji,A., Hine,C., Tanai,Y., Yonemoto,K., Mori,K., Yoshida,S.,
Bando,M., Sakai,E., Mori,K., Akamatsu,T. and Matsuda,Y.
TITLE Genomic organization and alternative splicing of human PACE4
(SPC4), kexin-like processing endoprotease
JOURNAL J. Biochem. 122 (2), 438-452 (1997)
MEDLINE 98021085
FEATURES Location/Qualifiers
source
1..3227
/organism="Homo sapiens"
/db_xref="taxon:9606"
join(AB001899.1:1158..1575,AB001900.1:906..908,
AB001900.1:1005..1010,AB001901.1:800..910,
AB001901.1:1434..1577,AB001901.1:2833..2909,
AB001901.1:4912..5000,AB001902.1:66..238,
AB001903.1:1698..1910,AB001904.1:292..392,
AB001905.1:819..922,AB001906.1:382..499,
AB001907.1:7012..7200,1154..1290,2441..2608)
/gene="PACE4"
/note="alternative splicing"
/product="PACE4"
join(AB001901.1:899..910,AB001901.1:1434..1577,
AB001901.1:2833..2909,AB001901.1:4912..5000,
AB001902.1:66..238,AB001903.1:1698..1910,
AB001904.1:292..392,AB001905.1:819..922,
AB001906.1:382..499,AB001907.1:7012..7200,1154..1290,
2441..2577)
/gene="PACE4"
/note="alternative splicing"
/codon_start=1
/product="PACE4"
/protein_id="BA21622.1"
/db_xref="GI:2281773"
/translaton="MYLHCGDKNSRCSRMNQAAWKRYTGKNNVVVITLDDGIERN
HPLAPNDYSYSDYNGNDYDPSPRYDASNNENKRGTRCAGEVAASANNISYCIYGLAY
NARIKGRLMDGVTDVWEAKSLGIRPNYIDYASWGDDDKTVDGPRLAQKQAF
YKIKGROGLGSLFVWASGNGREGDYCSODGYNTSYTSVSSATENGKYPWLEEC
ASTLATYSSGAEYERKIYVTDLRQCTDGHGTSTVSAPWAGIILALEANSQLTWR
DVUHLVKTSPRALHAKASDKWNGAGKVSHFYGLVDRAELVWEAKKWTAVPSQHM
CVAASDKRPSIPLVQVTRTALTSAHSDDRVLYLHVHVVRTSISHPRGDLQIY
LVSPGSKTQSLAKRLDLDSNEGTFNWFTHCWEKAEGQWTLIEIQLPQSQRNPE
KQDGLTPVANQLTTEERFVSTLSILFHSVYLSWSQYHIVLITVAL"
exon
1154..1290
/gene="PACE4"
/number=14
exon
2441..2608
/gene="PACE4"
/note="alternative splicing; exon 15CD"
/number=15
exon
2441..2608
/gene="PACE4"
/note="alternative splicing; exon 15D"
/number=15

```

```

BASE COUNT 799 a 690 c 678 g 1049 t 11 others
ORIGIN
..
alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-7 x AB00189511/rev ..
Align seg 1/1 to reverse of: AB00189511 from: 1 to: 3227
1 TyrArgLeuAlaIleArg 6
|||||
1773 TATCGTTTAGCCACGA 1756
seq_name: gb_pl1:YS2BGLUB
seq_documentation_block:
LOCUS YS2BGLUB 3371 bp DNA PLN 27-APR-1993
DEFINITION S.fibuligera extracellular beta-glucosidase 2 (BGL2) gene, complete
cds.
ACCESSION M22476
VERSION M22476.1 GI:170809
KEYWORDS BGL2 protein; beta-glucosidase 2.
SOURCE S.fibuligera (strain HUY7212) DNA, clones psfbetaG[1,2,3,4].
ORGANISM Saccharomycopsis fibuligera
Eukaryota; Fungi; Ascomycota; Saccharomycetales;
Saccharomycopsidaceae; Saccharomycopsis.
REFERENCE 1 (bases 1 to 3371)
AUTHORS Machida,M., Ohtsuki,I., Fukui,S. and Yamashita,I.
TITLE Nucleotide sequences of Saccharomycopsis fibuligera genes for
extracellular beta-glucosidases as expressed in Saccharomyces
cerevisiae
JOURNAL Appl. Environ. Microbiol. 54, 3147-3155 (1988)
MEDLINE 89133518
COMMENT Draft entry and sequence for [1] kindly submitted by I.Yamashita,
09-FEB-1989.
FEATURES Location/Qualifiers
source
1..3371
/organism="Saccharomycopsis fibuligera"
/db_xref="taxon:4944"
422..472
/note="beta-glucosidase 2 signal peptide"
422..3064
/note="beta-glucosidase 2 precursor"
/codon_start=1
/protein_id="AAA34315.1"
/db_xref="GI:170810"
/translaton="MLLILELVLIIGLVALPVQTHNLTNDQGEDESSOWISPHY
PPQGRGQGVQWDATYAKALVDSQMTIVKRVNLTGTGQWQPCVGNVTSVPRGIP
NLCLDQPGVRLTDFSTGPGMATGATFNKDLFQHQALGHFNFSKGVHIALGPA
VPLGKARGGNFEAFSDPYLQIAATAIKQENNVMACVKHFIGNEQDIYRQ
SNKSYDVPDPATKESISANIPDRAMHLYLWPFADSIKAGVSMCSYNRVNTYSC
ENSYNTHLLKEELGFGQVSDMAQKSGAYSAISGLDMSMPGLGGLGWNTKSYWG
QNLTKAVNVTPIERLDDMATKILALYATNSFTKQPLNPFSSFTTKEYNEFEVD
KSPKYNVNHVDPNDFTEDALKVAEESIVLLKNEKNTLIPSNKVRKLLLSGIAA
GDPKGYECSQDQVGFEGWGSQGVYQYPTTPEEISANARKNMKQFDYIKES
FOLTQVTSVDAHMSIVVYSAVSGEGLIIDNKGDKNNVTLNHNSNLIKAVAEK
ANTVVITSTQGVDSFADHPNVAIVWAGPLGDSGTAIANILFGNANPSGHLPT
VAKSNDVPIVYVNPNGPEPDNTLAEDLLDLYRYFEKNIERYAFGYGLSYNEY
KYSNAKVSAAKKVDEELPQKLYLAEYSVNTKEINNPDFAFPNSNARRIOEFLPYL
DSNVLKDCNIEYPDGYSTEQRTPIQGGGLGNDALWEVAYKVEVDVQNLGNSDK
FVPLKHPEDKGFETPVQLRGFENKVELSPGKNTVEFELLRLDLSWDDTTRQSWIV
ESGTYEALIGVAVNDIKSTFTI"
mat_peptide
473..3061
/note="beta-glucosidase 2"
BASE COUNT 975 a 641 c 701 g 1054 t
ORIGIN 3 bp upstream of PvuII site.

```


THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-7 to: N_Geneseq_36:* out_format : pfs

Date: Feb 8, 2000 1:27 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=frame+ p2n.model -DEV=alp
-Q/cnrl1/USPTO.spool/US08653294/runat_04022000_160701_15807/app_query.fasta.1
-DB=N_Geneseq_36 -QFMT=fastcap -Surfix=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -CGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-GAPOP=6.000 -CGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT
-THREADS=1

Search information block:

Query: US-08-653-294-7

Query length: 6

Database: N_Geneseq_36:*

Database sequences: 311585

Database length: 125096042

Search time (sec): 590.520000

score_list:

| Sequence | Strd | Orig | ZScore | EScore | Len | Documentation |
|---------------------|------|-------|--------|---------|-------|---------------------------------|
| N_Geneseq_36:T59906 | - | 28.00 | 112.74 | 111.47 | 1004 | Yeast transcription regulatory |
| N_Geneseq_36:V15446 | + | 28.00 | 108.66 | 188.26 | 959 | Human gene fragment Q28B8 from |
| N_Geneseq_36:X36423 | + | 28.00 | 101.57 | 467.21 | 2199 | Human hergulin-like factor cc |
| N_Geneseq_36:Q11477 | - | 28.00 | 94.95 | 1.1e+03 | 4775 | Sequence of clone encoding ald |
| N_Geneseq_36:N60174 | - | 28.00 | 94.95 | 1.1e+03 | 4775 | Alcohol-oxidase. New DNA seq |
| N_Geneseq_36:V62113 | - | 28.00 | 89.33 | 2.2e+03 | 9217 | HSV-2 strain SB5 Contig ID 101 |
| N_Geneseq_36:V62154 | + | 28.00 | 84.67 | 4.1e+03 | 15899 | HSV-2 strain SB5 Contig ID 3 |
| N_Geneseq_36:V62175 | + | 28.00 | 84.20 | 4.3e+03 | 16812 | HSV-2 strain SB5 Contig ID 12 |
| N_Geneseq_36:V72754 | + | 27.00 | 118.19 | 55.45 | 186 | Human genome fragment (prefer |
| N_Geneseq_36:N51895 | + | 27.00 | 114.75 | 86.12 | 278 | Human secreted protein 5' EST S |
| N_Geneseq_36:N81494 | + | 27.00 | 110.64 | 145.95 | 450 | Sequence encoding new physiolog |
| N_Geneseq_36:N81929 | + | 27.00 | 110.41 | 150.22 | 462 | EST clone CH421. New polynucle |
| N_Geneseq_36:X20674 | - | 27.00 | 95.88 | 969.51 | 2535 | Polynucleotide sequence from t |
| N_Geneseq_36:V03311 | + | 27.00 | 89.23 | 2.3e+03 | 5520 | Thermococcus 9N-2 phosphatase |
| N_Geneseq_36:V52143 | - | 27.00 | 84.24 | 4.3e+03 | 9897 | Streptococcus pneumoniae genom |
| N_Geneseq_36:N91695 | - | 27.00 | 83.17 | 4.9e+03 | 11225 | Sequence of fowlpox virus (PF |
| N_Geneseq_36:X21124 | + | 26.00 | 102.32 | 424.47 | 706 | Polynucleotide sequence from t |
| N_Geneseq_36:V52048 | + | 26.00 | 99.33 | 622.20 | 1001 | Helicobacter polypeptide GHPO |
| N_Geneseq_36:V06504 | - | 26.00 | 97.98 | 740.22 | 1173 | CBEV spheroidin gene. Spheroid |
| N_Geneseq_36:X14777 | + | 26.00 | 96.23 | 926.65 | 1440 | H. pylori GHPO 710 gene. New |
| N_Geneseq_36:V72549 | + | 26.00 | 96.02 | 952.06 | 1476 | Vernonia galamensis fatty ac |
| N_Geneseq_36:V52610 | + | 26.00 | 93.99 | 1.2e+03 | 1872 | Nucleotide sequence of lepidot |
| N_Geneseq_36:Q46816 | + | 26.00 | 93.20 | 1.4e+03 | 2053 | DTA gene. Nucleotide sequence |
| N_Geneseq_36:V38373 | - | 26.00 | 92.95 | 1.4e+03 | 2115 | Beta(1 -> 4)-N-acetylglucosam |
| N_Geneseq_36:V38368 | - | 26.00 | 92.43 | 1.5e+03 | 2246 | Beta(1 -> 4)-N-acetylglucosam |
| N_Geneseq_36:N81546 | - | 26.00 | 86.72 | 3.1e+03 | 4382 | Bio A, Bio B and Bio D-encodin |
| N_Geneseq_36:V13279 | + | 26.00 | 80.17 | 7.2e+03 | 9444 | cDNA to genomic hepatitis C vi |
| N_Geneseq_36:V17789 | + | 26.00 | 78.89 | 8.5e+03 | 10968 | Tomato Prf genomic DNA. Isol |
| N_Geneseq_36:X20562 | + | 26.00 | 72.96 | 1.8e+04 | 21948 | Polynucleotide sequence from |
| N_Geneseq_36:Q68993 | + | 25.00 | 112.29 | 118.12 | 130 | Drosophila-12 cadherin-related |
| N_Geneseq_36:V03617 | + | 25.00 | 112.29 | 118.12 | 130 | Protocadherin clone DROSOPHILA |
| N_Geneseq_36:V90602 | + | 25.00 | 106.02 | 264.10 | 271 | Nucleotide sequence of clone Y1 |
| N_Geneseq_36:V20113 | + | 25.00 | 104.18 | 334.23 | 336 | Probe (14) for microbial genes |
| N_Geneseq_36:T67813 | - | 25.00 | 102.09 | 436.79 | 429 | H. pylori cytoplasmic protein C |
| N_Geneseq_36:T77493 | - | 25.00 | 102.09 | 436.79 | 429 | H. pylori cytoplasmic protein C |
| N_Geneseq_36:Q97420 | + | 25.00 | 101.40 | 477.10 | 465 | Human interleukin-1beta gene. D |
| N_Geneseq_36:T28365 | - | 25.00 | 96.14 | 936.83 | 861 | Bacterial antibiotic resistance |
| N_Geneseq_36:V90638 | - | 25.00 | 96.07 | 945.18 | 868 | Nucleotide sequence of clone Y1 |
| N_Geneseq_36:N81774 | + | 25.00 | 95.99 | 955.92 | 877 | DNA coding sequence for gentam |
| N_Geneseq_36:T68253 | + | 25.00 | 94.65 | 1.1e+03 | 1026 | H. pylori cytoplasmic protein |
| N_Geneseq_36:X13499 | - | 25.00 | 93.12 | 1.4e+03 | 1227 | Enterococcus faecalis genome d |

N_Geneseq_36:Q31435 - 25.00 92.99 1.4e+03 1246 ! Encodes a Hela cell sialylt
N_Geneseq_36:Q66890 - 25.00 92.99 1.4e+03 1246 ! Sialyltransferase. Proteins
N_Geneseq_36:X30558 - 25.00 92.55 1.5e+03 1311 ! H. pylori outer membrane pr
N_Geneseq_36:X06864 - 25.00 92.48 1.5e+03 1322 ! Australian banana cv. Willi

seq_name: N_Geneseq_36:T59906

seq_documentation_block:

ID T59906 standard; DNA; 1004 BP.

AC T59906; 1997 (first entry)

DT 04-JUN-1997 (first entry)

DE Yeast transcription regulatory factor SRB6 DNA.

KW Transcription regulatory factor; suppressor of RNA polymerase B;

KW SRB6; RNA polymerase II; holoenzyme; SWI/SNF; ss.

OS Saccharomyces cerevisiae.

PH Key Location/Qualifiers

FT CDS 286..651

FT /*tag= a

PN WO9708301-A1.

PD 06-MAR-1997.

PF 28-AUG-1996; U14192.

PR 31-AUG-1995; US-521872.

PR 11-OCT-1995; US-540804.

PR 26-JAN-1996; US-590399.

PA (WHEED) WHITEHEAD INST BIOMEDICAL RES.

PI Chao DM, Koleske AJ, Thompson CM, Young RA;

DR WPI; 97-179258/16.

DR P-PSDB; W13823.

PT Purified RNA polymerase II holoenzyme - comprises RNA polymerase II

and one or more regulatory proteins, pref. suppressor of RNA

PT polymerase B proteins or SWI/SNF proteins

PS Claim 11; Fig 2C: 154pp; English.

CC Novel DNA sequences (T59904-11) code for yeast SRB (suppressor of

RNA polymerase B) proteins SRB4, SRB5, SRB6, SRB7, SRB8, SRB9,

CC SRB10 and SRB11 (W13821-28), respectively. SRBs are transcription

CC regulatory factors that act as positive and negative regulators

CC of RNA polymerase II activity. They are components of the RNA

CC polymerase II holoenzyme. The sequences were identified using

CC methods designed to identify transcription factors involved in RNA

CC polymerase II C-terminal domain (CTD) function. A genomic clone

CC for SRB6 was isolated by taking advantage of its ability to

CC suppress dominantly the cold-sensitive phenotype of a cell contg. a

CC CUPD. SRB nucleic acids and RNA polymerase II holoenzymes can be

CC used to treat diseases resulting from alteration or deletion of the

CC SRB gene, pref. by gene transfer technology. Probes based on the

CC genes can be used to detect SRB nucleic acids. 281 T;

SQ Sequence 1004 BP; 328 A; 193 C; 202 G;

alignment_scores:

Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-7 x T59906/rev ...

Align seg 1/1 to reverse of: T59906 from: 1 to: 1004

1 TyrArgLeuAlaIleArg 6

995 TATAGACTAGCAATAGG 978

seq_name: N_Geneseq_36:V15446

seq_documentation_block:

ID V15446 standard; DNA; 959 BP.

AC V15446;

DT 11-JUN-1998 (first entry)

DE Human gene fragment Q28B8 from chromosome 21q22.2 SEQ ID NO:22.

KW Human; chromosome 21; Down's syndrome; detection; isolation;

KW Carboxydehydrogenase; ds.

OS Homo sapiens.

PN J10057062-A.
 PD 03-MAR-1998.
 PF 16-AUG-1998; 216406.
 PR 16-AUG-1998; JP-216406.
 PA (RIKA) RIKAGAKU KENYUSHO.
 DR WPI; 98-254322/23.
 PT Detection and isolation of active gene - by physically mapping
 specific region on chromosome and detecting and isolating gene
 PS Example 1; Page 17; 21pp; Japanese.
 CC The present sequence represents a base sequence from the human
 chromosome 21, used in an example of the present invention.
 CC The present invention describes a method for detecting and
 isolating an active gene in which an uncommon restriction enzyme
 is prepared. The method comprises: (a) physically mapping a specified
 region of a chromosome and the base sequence upstream and downstream
 of the restriction enzyme site within the map, and (b) detecting and
 isolating the active gene.
 SQ Sequence 959 BP; 247 A; 240 C; 289 G; 183 T;

alignment_scores:
 Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
 US-08-653-294-7 x V15446 ..

Align seg 1/1 to: V15446 from: 1 to: 959

1 TyrArgLeuAlaileArg 6
 |||||
 594 TACAGGTGGCGTCAGA 611

seq_name: N_Geneseq_36:X36423

seq_documentation_block:
 ID X36423 standard; DNA; 2199 BP.
 AC X36423;
 DT 06-JUL-1999 (first entry)
 DE Human heregulin-like factor coding sequence.
 KW Human heregulin-like factor; HLF; cell growth regulator; diagnosis;
 OS neural system disorder; cancer; ss.
 HO Homo sapiens.
 PN WO9857989-A1.
 PD 23-DEC-1998.
 PF 16-JUN-1998; U12403.
 PR 17-JUN-1997; US-049942.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (GEOU) UNIV GEORGETOWN.
 PI Hijazi MM, King CR, Ruben SM, Young P;
 DR WPI; 99-095327/08.
 DR P-PSDB; Y05451.
 PT New isolated heregulin-like factor - used to develop products for
 the diagnosis and treatment of disorders involving regulation of
 cell growth, particularly cancers
 PS Claim 2; Page 86-87; 118pp; English.
 CC This sequence encodes the human heregulin-like factor (HLF) of the
 invention. The HLF is involved in the regulation of cell growth.
 CC Detection of different levels of expression of the HLF gene can be used
 for the diagnosis of disorders, e.g. in the neural system. In
 CC particular, detection of different levels of HLF gene expression in cells
 or body fluid of an individual can be used for diagnosing cancer. The
 CC products can also be used in the treatment of disorders involving
 abnormal levels of HLF activity.
 SQ Sequence 2199 BP; 689 A; 445 C; 462 G; 603 T;

alignment_scores:
 Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
 US-08-653-294-7 x X36423 ..

Align seg 1/1 to: X36423 from: 1 to: 2199

1 TyrArgLeuAlaileArg 6
 |||||
 1176 TATAGATTAGCTGTAGG 1193

seq_name: N_Geneseq_36:Q11477

seq_documentation_block:
 ID Q11477 standard; DNA; 4775 BP.
 AC Q11477;
 DT 20-JUN-1991 (first entry)
 DE Sequence of clone encoding alcohol oxidase, MOX enzyme.
 KW Oxidoreductase; bleaching; detergents; dihydroxyacetone synthase;
 PN alcohol oxidase; MOX; ss.
 EP-423890-A.
 PD 24-APR-1991.
 PF 25-JUL-1985; 202731.
 PR 27-JUL-1984; EP-201114.
 PR 07-FEB-1985; GB-003160.
 PR 01-JAN-1990; EP-202731.
 PA (UNIL) UNILEVER NV.
 PI Ledebor AM, Maat J, Verrips CT, Visser C, Janowicz ZA;
 PI Hollenberg CP;
 DR WPI; 91-119160/17.
 DR P-PSDB; R11606.
 PT Oxido-reductase prepn. - for use in bleaching and/or detergent
 compns. by recombinant DNA technology
 PS Disclosure; fig 11; 54pp; English.
 CC This sequence is contained in a plasmid along with regulatory
 DNA, e.g. a sequence allowing autonomous replication in host
 cells, for recombinant prodn. of MOX in high yields. MOX catal-
 yses the oxidation of methanol to formaldehyde and hydrogen
 peroxide in the bleaching process and is useful in bleach and/or
 detergent compns.
 CC See also Q11476 and Q11478-81.
 SQ Sequence 4775 BP; 1171 A; 1249 C; 1286 G; 1069 T;

alignment_scores:
 Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
 US-08-653-294-7 x Q11477/rev ..

Align seg 1/1 to reverse of: Q11477 from: 1 to: 4775

1 TyrArgLeuAlaileArg 6
 |||||
 3055 TACAGGTGGCAGTGAGG 3038

seq_name: N_Geneseq_36:N60174

seq_documentation_block:
 ID N60174 standard; DNA; 4775 BP.
 AC N60174;
 DT 01-AUG-1991 (first entry)
 DE Alcohol-oxidase.
 KW Alcohol-oxidase; ss.
 FH Key Location/Qualifiers
 FT cds 1512..3504
 FT /*tag= a
 FT /label= alcohol-oxidase gene
 PN EP-173378-A.
 PD 05-MAR-1986.
 PF 25-JUL-1985; 201235.
 PR 27-JUL-1984; EP-201114.
 PR 07-FEB-1985; GB-003160.

| | | |
|----|--|--|
| FT | /*tag= h | |
| FT | /product= "ORF#8 protein" | |
| FT | /note= "encoded protein shown in W72014" | |
| PN | WO9820016-A1. | |
| PD | 14-MAY-1998. | |
| PD | U20016. | |
| PP | 31-OCT-1997; | |
| PP | 09-JUN-1997; US-049018. | |
| PP | 04-NOV-1996; US-030279. | |
| PA | (SMIK) SMITHKLINE BEECHAM CORP. | |
| PA | Chan JY, dabrowski-Amaral CE, Delvecchio AM, Dillon SB, | |
| PI | Esser KM, Leary JJ; | |
| PI | WPI: 98-296847/25; | |
| DR | P-SDB; W72007, W72008, W72009, W72010, W72011, W72012. | |
| PT | Herpes simplex virus type-2 sequences- useful in, e.g. | |
| PT | and treatment of infection or inducing immunological res | |
| PT | manual | |
| PT | Claim 1; Page 175-179; 748pp; English. | |
| CC | This sequence represents a Herpes simplex virus type-2 (| |

sequence of the invention. This sequence was isolated from Hsv-2 strain SB5 (deposited as ATCC VR-2346), is designated Conlig ID 101, and encodes 8 Hsv-2 proteins. The proteins can be used for the treatment or prevention of disease, to induce an immunological response in a mammal or to identify inhibitors, activators or novel antivirals. Antagonists of the proteins can be used to inhibit a viral polypeptide. The DNA sequence or a vector containing it can also be used to induce an immunological response in a mammal.

Sequence 9217 BP: 1594 A: 2837 C: 3250 G: 1527 T:

CC the proteins can be used to inhibit a viral protease, the DNA sequence
CC or a vector containing it can also be used to induce an immunological
CC response in a mammal.

| | | | |
|----|-----------------------|---------------------------------|--|
| CC | response in a mammal. | | |
| SQ | Sequence | 1594 A; 2837 C; 3250 G; 1527 T; | |

sequence 341/ 342, 343/ 344, 345/ 346, 347/ 348, 349/ 350, 351/ 352, 353/ 354, 355/ 356, 357/ 358, 359/ 360, 361/ 362, 363/ 364, 365/ 366, 367/ 368, 369/ 370, 371/ 372, 373/ 374, 375/ 376, 377/ 378, 379/ 380, 381/ 382, 383/ 384, 385/ 386, 387/ 388, 389/ 390, 391/ 392, 393/ 394, 395/ 396, 397/ 398, 399/ 400, 401/ 402, 403/ 404, 405/ 406, 407/ 408, 409/ 410, 411/ 412, 413/ 414, 415/ 416, 417/ 418, 419/ 420, 421/ 422, 423/ 424, 425/ 426, 427/ 428, 429/ 430, 431/ 432, 433/ 434, 435/ 436, 437/ 438, 439/ 440, 441/ 442, 443/ 444, 445/ 446, 447/ 448, 449/ 450, 451/ 452, 453/ 454, 455/ 456, 457/ 458, 459/ 460, 461/ 462, 463/ 464, 465/ 466, 467/ 468, 469/ 470, 471/ 472, 473/ 474, 475/ 476, 477/ 478, 479/ 480, 481/ 482, 483/ 484, 485/ 486, 487/ 488, 489/ 490, 491/ 492, 493/ 494, 495/ 496, 497/ 498, 499/ 500, 501/ 502, 503/ 504, 505/ 506, 507/ 508, 509/ 510, 511/ 512, 513/ 514, 515/ 516, 517/ 518, 519/ 520, 521/ 522, 523/ 524, 525/ 526, 527/ 528, 529/ 530, 531/ 532, 533/ 534, 535/ 536, 537/ 538, 539/ 540, 541/ 542, 543/ 544, 545/ 546, 547/ 548, 549/ 550, 551/ 552, 553/ 554, 555/ 556, 557/ 558, 559/ 560, 561/ 562, 563/ 564, 565/ 566, 567/ 568, 569/ 570, 571/ 572, 573/ 574, 575/ 576, 577/ 578, 579/ 580, 581/ 582, 583/ 584, 585/ 586, 587/ 588, 589/ 590, 591/ 592, 593/ 594, 595/ 596, 597/ 598, 599/ 600, 601/ 602, 603/ 604, 605/ 606, 607/ 608, 609/ 610, 611/ 612, 613/ 614, 615/ 616, 617/ 618, 619/ 620, 621/ 622, 623/ 624, 625/ 626, 627/ 628, 629/ 630, 631/ 632, 633/ 634, 635/ 636, 637/ 638, 639/ 640, 641/ 642, 643/ 644, 645/ 646, 647/ 648, 649/ 650, 651/ 652, 653/ 654, 655/ 656, 657/ 658, 659/ 660, 661/ 662, 663/ 664, 665/ 666, 667/ 668, 669/ 670, 671/ 672, 673/ 674, 675/ 676, 677/ 678, 679/ 680, 681/ 682, 683/ 684, 685/ 686, 687/ 688, 689/ 690, 691/ 692, 693/ 694, 695/ 696, 697/ 698, 699/ 700, 701/ 702, 703/ 704, 705/ 706, 707/ 708, 709/ 710, 711/ 712, 713/ 714, 715/ 716, 717/ 718, 719/ 720, 721/ 722, 723/ 724, 725/ 726, 727/ 728, 729/ 730, 731/ 732, 733/ 734, 735/ 736, 737/ 738, 739/ 740, 741/ 742, 743/ 744, 745/ 746, 747/ 748, 749/ 750, 751/ 752, 753/ 754, 755/ 756, 757/ 758, 759/ 760, 761/ 762, 763/ 764, 765/ 766, 767/ 768, 769/ 770, 771/ 772, 773/ 774, 775/ 776, 777/ 778, 779/ 780, 781/ 782, 783/ 784, 785/ 786, 787/ 788, 789/ 790, 791/ 792, 793/ 794, 795/ 796, 797/ 798, 799/ 800, 801/ 802, 803/ 804, 805/ 806, 807/ 808, 809/ 810, 811/ 812, 813/ 814, 815/ 816, 817/ 818, 819/ 820, 821/ 822, 823/ 824, 825/ 826, 827/ 828, 829/ 830, 831/ 832, 833/ 834, 835/ 836, 837/ 838, 839/ 840, 841/ 842, 843/ 844, 845/ 846, 847/ 848, 849/ 850, 851/ 852, 853/ 854, 855/ 856, 857/ 858, 859/ 860, 861/ 862, 863/ 864, 865/ 866, 867/ 868, 869/ 870, 871/ 872, 873/ 874, 875/ 876, 877/ 878, 879/ 880, 881/ 882, 883/ 884, 885/ 886, 887/ 888, 889/ 890, 891/ 892, 893/ 894, 895/ 896, 897/ 898, 899/ 900, 901/ 902, 903/ 904, 905/ 906, 907/ 908, 909/ 910, 911/ 912, 913/ 914, 915/ 916, 917/ 918, 919/ 920, 921/ 922, 923/ 924, 925/ 926, 927/ 928, 929/ 930, 931/ 932, 933/ 934, 935/ 936, 937/ 938, 939/ 940, 941/ 942, 943/ 944, 945/ 946, 947/ 948, 949/ 950, 951/ 952, 953/ 954, 955/ 956, 957/ 958, 959/ 960, 961/ 962, 963/ 964, 965/ 966, 967/ 968, 969/ 970, 971/ 972, 973/ 974, 975/ 976, 977/ 978, 979/ 980, 981/ 982, 983/ 984, 985/ 986, 987/ 988, 989/ 990, 991/ 992, 993/ 994, 995/ 996, 997/ 998, 999/ 1000, 1001/ 1002, 1003/ 1004, 1005/ 1006, 1007/ 1008, 1009/ 1010, 1011/ 1012, 1013/ 1014, 1015/ 1016, 1017/ 1018, 1019/ 1020, 1021/ 1022, 1023/ 1024, 1025/ 1026, 1027/ 1028, 1029/ 1030, 1031/ 1032, 1033/ 1034, 1035/ 1036, 1037/ 1038, 1039/ 1040, 1041/ 1042, 1043/ 1044, 1045/ 1046, 1047/ 1048, 1049/ 1050, 1051/ 1052, 1053/ 1054, 1055/ 1056, 1057/ 1058, 1059/ 1060, 1061/ 1062, 1063/ 1064, 1065/ 1066, 1067/ 1068, 1069/ 1070, 1071/ 1072, 1073/ 1074, 1075/ 1076, 1077/ 1078, 1079/ 1080, 1081/ 1082, 1083/ 1084, 1085/ 1086, 1087/ 1088, 1089/ 1090, 1091/ 1092, 1093/ 1094, 1095/ 1096, 1097/ 1098, 1099/ 1100, 1101/ 1102, 1103/ 1104, 1105/ 1106, 1107/ 1108, 1109/ 1110, 1111/ 1112, 1113/ 1114, 1115/ 1116, 1117/ 1118, 1119/ 1120, 1121/ 1122, 1123/ 1124, 1125/ 1126, 1127/ 1128, 1129/ 1130, 1131/ 1132,

| alignment_scores: | |
|-------------------|--------|
| Quality | length |
| 28 | 6 |

| | | | |
|---------------------|---------|-------------------|--------|
| Percent Similarity: | 100.000 | Percent Identity: | 93.333 |
| Ratio: | 4.667 | Gaps: | 0 |
| Quality: | 28.000 | Length: | 0 |

Percent Similarity: 100.000 Percent Identity: 83.333

US-08-653-294-7 x V62131/rev ..

Align seq 1/1 to reverse of: v62131 from: 1 to: 9217

Align seg 1/1 to reverse of: V62131 from: 1 to: 9217

seq name: N Geneseq 36:V62164

seq_name: N_Geneseq_36:V62164

```
seq documentation block:
```

seq_documentation_block:
ID V62164 standard; DNA; 15899 BP.

| ID | V62164 standard; DNA; 15899 BP. |
|----|---------------------------------|
| AC | V62164; |

| | |
|----|---------------------------|
| AC | V62164; |
| DT | 23-DEC-1998 (first entry) |

DT 23-DEC-1998 (first entry)
DE HSV-2 strain SB5 Contig ID 3 DNA sequence.

DE
HSV-2 strain SB5 Contlig ID 3 DNA sequence.
KW
HSV-2 strain SB5; immunological response induction; therapy;
antigenic identification; viral protein inhibitors; co

NSV-2 strain SBJ; immunological response induction; therapy;
antiviral identification; viral protein inhibitor; ss.
Hermes simplex virus type 2
Hermes simplex virus type 2

| | |
|----|---|
| OS | Herpes simplex virus type 2. |
| OS | antiviral identification; viral protein inhibitors; SS: |
| FH | Key Location/Qualifiers |

| | | |
|----|-----|---------------------|
| FH | Key | Location/Qualifiers |
| FT | CDS | 1531. .2775 |

FT CDS 1531.2775
FT /tag= a

```
ET /tag= a
ET /product= "ORF#1 protein"
```

| Accession | Protein | Product | Note |
|-----------|---------------------------------|--|------|
| FT | ORF#1 protein | /product= "ORF#1 protein" | |
| FT | encoded protein shown in W72139 | /note= "encoded protein shown in W72139" | |
| FT | | | |
| FT | | | |

| | |
|----|--|
| FT | /note= "encoded protein shown in w/4139" |
| FT | complement (2957. .3832) |
| FT | /*tag= b |

```

CDS
FT             complement(2557..10632)
FT             /tag=b
FT             /product="OBP#2 protein"

```

```
/product= "ORF#2 protein"  
/note= "encoded protein shown in W72140"
```

| | |
|----|--|
| | CDS |
| FT | /note= "encoded protein shown in W72140" |
| FT | 4028..5563 |

```
FT CDS
4028.5563
/*tag= c
```

```
FT      /*tag= C
FT      /product= "ORE#3 protein"
```

```

FT
/product="ORF#3 protein"
FT
/note="encoded protein shown in W72141"

```

| FT | CDS | /note= "encoded protein shown in W2141" |
|----|------------|---|
| FT | 8239..8352 | |
| FT | 8353..8353 | |
| FT | 8354..8354 | |

```
ET CDS
      8239. .8392
      /*tag= d
```

```

z1      /tag= u
PT      /product= "ORF#4 protein"
PT      /note= "encoded protein shown in W721A2"

```

| | |
|----|--|
| FT | /product= ONK4 protein |
| FT | /note= "encoded protein shown in W72142" |
| FT | 8775 9977 |
| FT | CDS |

```
FT CDS  
      /REC= ONCE PER YEAR IN MARCH  
FT     /TAG= P
```

```
ET /*tag= e
ET /product= "ORF#5 protein"
```

FT /product= "ORF#5 protein"

FT CDS /note= "encoded protein shown in W72143"
 9396. .11297
 FT /tag= f
 FT /product= "ORF#6 protein"
 FT /note= "encoded protein shown in W72144"
 11616. .12968
 FT /tag= g
 FT /product= "ORF#7 protein"
 FT /note= "encoded protein shown in W72145"
 13415. .13552
 FT /tag= h
 FT /product= "ORF#8 protein"
 FT /note= "encoded protein shown in W72146"
 13652. .13921
 FT /tag= i
 FT /product= "ORF#9 protein"
 FT /note= "encoded protein shown in W72147"
 complement (14409. .15317)
 FT /tag= j
 FT /product= "ORF#10 protein"
 FT /note= "encoded protein shown in W72148"

PN W09820016-A1.

PD 14-MAY-1998.
 PF 31-OCT-1997; U20016.
 PR 09-JUN-1997; US-049018.
 PR 04-NOV-1996; US-030279.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Chan JY, Dabrowski-Amarel CE, Delvecchio AM, Dillon SB,
 PI Esser KM, Leary JJ;
 DR WPI: 98-286847/25.
 DR P-PSDB: W72139, W72140, W72141, W72142, W72143, W72144, W72145, W72146,
 DR W72147, W72148.
 PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention
 PT and treatment of infection or inducing immunological response in
 PT mammal
 PS Claim 1; Page 466-472; 748pp; English.
 CC This sequence represents a Herpes simplex virus type-2 (HSV-2) DNA
 CC sequence of the invention. This sequence was isolated from HSV-2 strain
 CC SB5 (deposited as ATCC VR-2546), is designated Contig ID 2, and encodes
 CC 10 HSV-2 proteins. The proteins can be used for the treatment or
 CC prevention of disease, to induce an immunological response in
 CC to identify inhibitors, activators or novel antivirals. Antagonists of
 CC the proteins can be used to inhibit a viral polypeptide. The DNA sequence
 CC or a vector containing it can also be used to induce an immunological
 CC response in a mammal.
 SQ Sequence 15899 BP; 2599 A; 5470 C; 5204 G; 2626 T;

alignment_scores:
 Quality: 28.00 Length: 6
 Ratio: 4.667 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:

US-08-653-294-7 x V62164 ..

Align seg 1/1 to: V62164 from: 1 to: 15899

1 TyrArgLeuAlaIleArg 6

12519 TACCGCTGGCGGTCGC 12536

seq_name: N_Geneseq_36.V62175

seq_documentation_block:

ID V62175 standard; DNA; 16812 BP.

AC V62175;

DT 08-JAN-1999 (first entry)

DE HSV-2 strain SB5 Contig ID 12 DNA sequence.

KW HSV-2 strain SB5; immunological response induction; therapy;

OS antiviral identification; viral protein inhibitor; ss.

OS Herpes simplex virus type 2.

Key Location/Qualifiers

FT CDS 127. .1371
 FT /tag= a
 FT /product= "ORF#1 protein"
 FT /note= "encoded protein shown in W72159"
 complement (1553. .2428)
 FT /tag= b
 FT /product= "ORF#2 protein"
 FT /note= "encoded protein shown in W72160"
 2714. .4159
 FT /tag= c
 FT /product= "ORF#3 protein"
 FT /note= "encoded protein shown in W72161"
 6835. .6948
 FT /tag= d
 FT /product= "ORF#4 protein"
 FT /note= "encoded protein shown in W72162"
 7392. .8573
 FT /tag= e
 FT /product= "ORF#5 protein"
 FT /note= "encoded protein shown in W72163"
 8775. .9893
 FT /tag= f
 FT /product= "ORF#6 protein"
 FT /note= "encoded protein shown in W72164"
 10212. .11858
 FT /tag= g
 FT /product= "ORF#7 protein"
 FT /note= "encoded protein shown in W72165"
 12010. .12147
 FT /tag= h
 FT /product= "ORF#8 protein"
 FT /note= "encoded protein shown in W72166"
 12247. .12516
 FT /tag= i
 FT /product= "ORF#9 protein"
 FT /note= "encoded protein shown in W72167"
 complement (13004. .13912)
 FT /tag= j
 FT /product= "ORF#10 protein"
 FT /note= "encoded protein shown in W72168"
 15899. .16582
 FT /tag= k
 FT /product= "ORF#11 protein"
 FT /note= "encoded protein shown in W72169"
 W09820016-A1.
 PN 14-MAY-1998.
 PF 31-OCT-1997; U20016.
 PR 09-JUN-1997; US-049018.
 PR 04-NOV-1996; US-030279.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Chan JY, Dabrowski-Amarel CE, Delvecchio AM, Dillon SB,
 PI Esser KM, Leary JJ;
 DR WPI: 98-286847/25.
 DR P-PSDB: W72159, W72160, W72161, W72162, W72163, W72164, W72165, W72166,
 DR W72167, W72168, W72169.
 PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention
 PT and treatment of infection or inducing immunological response in
 PT mammal
 PS Claim 1; Page 505-512; 748pp; English.
 CC This sequence represents a Herpes simplex virus type-2 (HSV-2) DNA
 CC sequence of the invention. This sequence was isolated from HSV-2 strain
 CC SB5 (deposited as ATCC VR-2546), is designated Contig ID 12, and encodes
 CC 11 HSV-2 proteins. The proteins can be used for the treatment or
 CC prevention of disease, to induce an immunological response in a mammal or
 CC to identify inhibitors, activators or novel antivirals. Antagonists of
 CC the proteins can be used to inhibit a viral polypeptide. The DNA sequence
 CC or a vector containing it can also be used to induce an immunological
 CC response in a mammal.
 SQ Sequence 16812 BP; 2708 A; 5989 C; 5367 G; 2748 T;

alignment_scores:

Quality: 28.00

Length: 6

Ratio: 4.667 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:

US-08-653-294-7 x V62175 ..

Align seg 1/1 to: V62175 from: 1 to: 16812

1 TyrArgLeuAlaIleArg 6

|||||.....|

11115 TACCCGCTGGCGTCCGC 11132

seq_name: N_Geneseq_36:Q77254

seq_documentation_block:

ID Q77254 standard; DNA; 186 BP.

AC Q77254;

DT 23-SEP-1994 (first entry)

DE Human genome fragment (Preferred).

KW Brain; placenta; bone marrow; genetic analysis; gene mapping;

KW detection; homology; human; adrenal tissue; ds.

OS Homo sapiens.

PN WO9401548-A.

PD 20-JAN-1994.

PF 13-JUL-1993; G01467.

PR 13-JUL-1993; GB-014857.

PA (MED1-) MEDICAL RES COUNCIL.

PI Gross J, Hadfield KM, Howells D, Kelly M, Shaw D;

PI Sibson DR, Starkey M;

DR WPI; 94-030506/04.

PT New nucleic acid fragment encoding gene products - can be used

PT for genetic analysis and mapping

PS Claim 1; Page 430; 616pp; English.

CC Human nucleic acid fragments, isolated from brain, adrenal tissue,

CC the placenta or bone marrow comprise any of: (A) a sequence

CC selected from (Q76401-Q77613), (B) an allelic variation of a

CC sequence as described in (A), or (C) a sequence complementary

CC to (A) or (B).

CC Preferred sequences exhibit no more than 90% homology to a human

CC sequence known per se.

SO Sequence 186 BP; 46 A; 36 C; 47 G; 57 T;

alignment_scores:

Quality: 27.00 Length: 6

Ratio: 4.500 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:

US-08-653-294-7 x Q77254 ..

Align seg 1/1 to: Q77254 from: 1 to: 186

1 TyrArgLeuAlaIleArg 6

|||||.....|

138 TATAGGATGGCATTAGG 155

seq_name: N_Geneseq_36:X51895

seq_documentation_block:

ID X51895 standard; DNA; 278 BP.

AC X51895;

DT 22-JUN-1999 (first entry)

DE Human secreted protein 5' EST SEQ ID NO: 109.

KW Human; secreted protein; EST; expressed sequence tag; diagnosis;

KW forensic; gene therapy; chromosome mapping; signal peptide;

KW upstream regulatory sequence; cytokine activity; cell proliferation;

KW differentiation; haematopoiesis regulation; tissue growth regulation;

KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;

KW thrombolytic; anti-inflammatory; tumour inhibition; ds.

OS Homo sapiens.

PN WO9906552-A2.

PD 11-FEB-1999.

PF 31-JUL-1998; IB1236.
PR 01-AUG-1997; US-905223.
PA (GEST) GENSET.
PI Ducicert A, Dumas Milne Edwards J, Lacroix B;
DR WPI; 99-153782/13.
DR P-PSDB; Y13095.
PT New isolated brain-derived nucleic acids - used to develop products
PT which may have cytokine, immune, regulatory, haematopoiesis
PT regulating, anti-inflammatory or tumour inhibition activity
PS Claim 1; Page 256; 577pp; English.
CC X51787 to X52019 represent 5' expressed sequence tags (ESTs) for human
CC secreted proteins, and encode the proteins given in Y12987 to Y13219,
CC respectively. The proteins given represent the signal peptide and an
CC N-terminal fragment of a secreted protein. The nucleic acid sequences
CC can be used for producing secreted human gene products. They can also
CC be used to develop products for diagnosis and therapy. The proteins
CC obtained may have cytokine activity, cell proliferation/differentiation
CC activity, haematopoiesis regulating activity, tissue growth regulating
CC activity, reproductive hormone regulating activity, chemotactic/
CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC ligand activity, anti-inflammatory activity, tumour inhibition activity
CC or other activities. The products can be used in forensic, gene therapy
CC and chromosome mapping procedures. The sequences can also be used for
CC obtaining corresponding promoter sequences. The nucleic acids encoding
CC the signal peptide can be used for directing extracellular secretion of
CC a polypeptide or the insertion of a polypeptide into a membrane, or
CC importing a polypeptide into a cell.
SO Sequence 278 BP; 95 A; 56 C; 58 G; 67 T;

alignment_scores:

Quality: 27.00 Length: 6

Ratio: 4.500 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:

US-08-653-294-7 x X51895 ..

Align seg 1/1 to: X51895 from: 1 to: 278

1 TyrArgLeuAlaIleArg 6

|||||.....|

241 TACAGACTAGCTCTGCGC 258

seq_name: N_Geneseq_36:N81494

seq_documentation_block:

ID N81494 standard; DNA; 450 BP.

AC N81494;

DT 14-SEP-1990 (first entry)

DE Sequence encoding new physiologically active polypeptide with antitumour

DE activity

FW Antitumour; Escherichia coli; ss.

FT Key Location/Qualifiers

mat_peptide 1..450

FT /*tag= a

PN J63226298-A.

PD 20-SEP-1988.

PF 16-MAR-1987; 059007.

PR 16-MAR-1987; JP-059007.

PA (TEIU) Teijin Kk.

PI

DR WPI; 88-305170/43.

DR P-PSDB; P81068.

PT New physiologically active polypeptide used as antitumour agent -

PT obtd. by culturing recombinant Escherichia coli cell transformed by

PT plasmid contg. DNA region coding the polypeptide

PS Claim 2 (4); Page 670; 20pp; Japanese

CC A recombinant microorganism cell transformed by recombinant plasmid

CC contg. the DNA region coding for the polypeptide is cultured. The

CC polypeptide is isolated from the resulting culture. The microorganism

CC is pref. Escherichia coli.

SO Sequence 450 BP; 95 A; 143 C; 126 G; 86 T;

```

alignment_scores:
  Quality: 27.00      Length: 6
  Ratio: 4.500        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 83.333

alignment_block:
  US-08-653-294-7 x N81494/rev ..
  Align seg 1/1 to reverse of: N81494 from: 1 to: 450

  1 TyxArgLeuAlaIleArg 6
  |||||
  18 TACAGGCTTCGGCTCGG 1

seq_name: N_Geneseq_36.V89229
seq_documentation_block:
  ID V89229 standard; CDNA; 462 BP.
  AC V89229;
  DT 15-FEB-1999 (first entry)
  DE EST clone CH421.
  KW Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
  KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
  KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
  KW gene therapy; ss.
  OS Homo sapiens.
  PN WO9845436-A2.
  PD 15-OCT-1998.
  PF 10-APR-1998: U06955.
  PR 10-APR-1997: US-838821.
  PA (GEMT) GENETICS INST INC.
  PI Agostino MJ, Jacobs K, Lavallee ER, McCoy JM, Merberg D,
  PI Racine LA, Spaulding V, Treacy M;
  DR WPI; 99-070077/06.
  PT New polynucleotides encoding human secreted proteins - derived from
  PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
  PT ovary, pituitary, retina and colon cDNA libraries.
  PS Claim 1; Page 151; 618pp; English.
  CC The present sequence represents a human expressed sequence tag (EST).
  CC The polynucleotide, which is a secreted EST, and the encoded protein
  CC are predicted to have useful biological activities which would make
  CC them suitable for treating, preventing or ameliorating medical
  CC conditions in humans and animals, although no supporting data is
  CC given. Suggested activities include nutritional activity, immune
  CC stimulating or suppressing activity, haematopoiesis regulating
  CC activity, tissue growth activity, activin/inhibin activity,
  CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
  CC activity, receptor/ligand activity, anti-inflammatory activity,
  CC cadherin/tumour invasion suppressor activity, tumour inhibition
  CC activity. The polynucleotide may also be useful for gene therapy.
  SQ Sequence 462 BP; 143 A; 81 C; 97 G; 141 T;

alignment_scores:
  Quality: 27.00      Length: 6
  Ratio: 4.500        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 83.333

alignment_block:
  US-08-653-294-7 x V89229 ..
  Align seg 1/1 to: V89229 from: 1 to: 462

  1 TyxArgLeuAlaIleArg 6
  |||||
  79 TACGCGCTAGCACTGAG 96

seq_name: N_Geneseq_36.X20674
seq_documentation_block:
  ID X20674 standard; DNA; 2535 BP.
  AC X20674;
  DT 05-MAY-1999 (first entry)
  DE Polynucleotide sequence from the genome of Treponema pallidum.
  KW Treponema pallidum infection; syphilis; Borrelia infection; animal;
  KW enzyme production; ds.
  OS Treponema pallidum.
  PN WO9859034-A2.
  PD 30-DEC-1998.
  PF 23-JUN-1998: U13041.
  PR 24-JUN-1997: US-050667.
  PA (HUMA-) HUMAN GENOME SCI INC.
  PI Fraser CM;
  DR WPI; 99-081273/07.
  PT New isolated Treponema pallidum nucleic acids - used to develop
  PT products for the detection, diagnosis, characterisation, prevention
  PT and therapy of T. pallidum infections, particularly syphilis
  PS Claim 1; Page 840-841; 1150pp; English.
  CC X20500-21243 represent polynucleotide sequences from the genome of
  CC Treponema pallidum. The sequences can be used for detection,
  CC diagnosis, characterisation, prevention and therapy for T. pallidum
  CC infections, particularly syphilis. They can also be used for detecting
  CC diseases related to Borrelia infections in animals, and for the
  CC production of biosynthetic products such as enzymes.
  SQ Sequence 2535 BP; 651 A; 831 C; 597 G; 455 T;

alignment_scores:
  Quality: 27.00      Length: 6
  Ratio: 4.500        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 83.333

alignment_block:
  US-08-653-294-7 x X20674/rev ..
  Align seg 1/1 to reverse of: X20674 from: 1 to: 2535

  1 TyxArgLeuAlaIleArg 6
  |||||
  327 TACGCGCTTCGGCTCGG 310

seq_name: N_Geneseq_36.V03311
seq_documentation_block:
  ID V03311 standard; DNA; 5520 BP.
  AC V03311;
  DT 22-JUN-1998 (first entry)
  DE Thermococcus 9N-2 phosphatase (21phl) gene.
  KW Alkaline phosphatase; thermostable enzyme; thermophilic bacterium;
  KW food; detergent; baking; ss.
  OS Thermococcus sp. strain 9N-2.
  PN WO9748416-A1.
  PD 24-DEC-1997.
  PF 19-JUN-1997: U10784.
  PR 19-JUN-1996; US-033752.
  PA (RECO-) RECOMBINANT BIOCATALYSIS INC.
  PI Bylina E, Lee E, Mathur EJ;
  DR WPI; 98-062851/06.
  PT Thermostable phosphatase(s) - useful in pharmaceutical, food,
  PT detergent, and baking industries
  PS Claim 12; Page 77-81; 128pp; English.
  CC This polynucleotide encodes a thermostable phosphatase, designated
  CC 31phl, of Thermococcus 9N-2. The invention relates to claimed
  CC polynucleotides (see V03301-20) coding for novel thermostable
  CC phosphatases (see W42380-95). Claimed vector and host cells are
  CC used to produce the enzymes, which can be used in a claimed method
  CC to hydrolyse phosphate bonds. They can also be used in enzyme
  CC labelling processes, in certain recombinant DNA techniques, in
  CC ELISA immunoassays, in enzyme linked gene probes, in research
  CC applications for removing 5' phosphates in polynucleotides prior to
  CC end labelling and in the pharmaceutical, food, detergent, and
  CC baking industries. Polynucleotides can also be used as probes.
  SQ Sequence 5520 BP; 1297 A; 1491 C; 1587 G; 1145 T;

```

1 TyrArgLeuAlaIleArg 6
|||||
8487 TACCGACTAGCTTTAAGG 8470

alignment_scores:
Quality: 27.00 Length: 6
Ratio: 4.500 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:

US-08-653-294-7 x V03311

Align seg 1/1 to: V03311 from: 1 to: 5520

1 TyrArgLeuAlaIleArg 6

|||||

2039 TACCGCTGCGCCTACGA 2056

seq_name: N_Geneseq_36:V52143

seq_documentation_block:

ID V52143 standard; DNA; 9897 BP.

AC V52143;

DT 23-OCT-1998 (first entry)

DE Streptococcus pneumoniae genome fragment SEQ ID NO:10.

KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;

KW computer readable medium; vaccine; pharmaceutical composition; ds.

OS Streptococcus pneumoniae.

PN W09818931-A2.

PD 07-MAY-1998.

PF 30-OCT-1997; U19588

PR 31-OCT-1996; US-029960.

PA (HUMAN) HUMAN GENOME SCI INC.

PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M,

PI Kunsch CA, Rosen CA;

DR WPI: 98-272225/24.

PT Computer-readable medium with recorded Streptococcus pneumoniae

PT polynucleotide sequences - useful in diagnostic kits and assays, and

PT pharmaceutical compositions and vaccines for Streptococcus

PT pneumoniae

PS Claim 1: Page 214-220; 1409pp; English.

CC The present invention describes a computer readable medium which has

CC the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded

CC on it, or a representative fragment or a sequence at least 95% identical

CC to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1

CC to 391 (V52134 to V52524) are genomic fragments from Streptococcus

CC pneumoniae. The present invention also describes an isolated nucleic acid

CC molecule encoding a homologue of any of the fragments of the S. pneumoniae

CC genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced

CC by a process comprising: (a) screening a genomic DNA library using as a

CC probe a target sequence defined by any of the sequences in SEQ ID NO:1

CC to 391, identifying members of the library which contain sequences

CC that hybridize to the target sequence and isolating the nucleic acid

CC molecules from the members; or (b) isolating mRNA, DNA or cDNA produced

CC from an organism, amplifying nucleic acid molecules whose nucleotide

CC sequence is homologous to amplification primers derived from the

CC fragment of the S. pneumoniae genome to prime the amplification and

CC isolating the amplified sequences. The computer readable medium can be

CC used in a computer-based system for identifying fragments of the

CC S. pneumoniae genome of commercial importance, or expression modulating

CC fragments of the S. pneumoniae genome. Products from the present

CC invention can be used in diagnosis kits and assays, and pharmaceutical

CC compositions and vaccines for S. pneumoniae.

SQ Sequence 9897 BP; 3110 A; 1980 C; 1493 G; 3312 T;

alignment_scores:

Quality: 27.00 Length: 6
Ratio: 4.500 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:

US-08-653-294-7 x V52143/rev

Align seg 1/1 to reverse of: V52143 from: 1 to: 9897

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-7 to: EST:* out_format : pfs
 Date: Feb 8, 2000 4:02 AM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 Compugen Ltd.

Command-line parameters:
 -MODEL=framet_p2n.model -DEV=xl
 -Q/cnrl1/uspto.spool/US08653294/runat_04022000_160700_15770/app_query.fasta.1
 -DB=EST -QFMT=fastp -SUFFIX=est -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -GAPOP=4.500
 -CGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -CGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -STARF=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -ALIGN=15 -MODE=LOCAL
 -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294
 -NGPU=6 -ICPU=3 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-7
 Query length: 6
 Database: EST:
 Database sequences: 4538634
 Database length: 1887831982
 Search time (sec): 8553.360000

| score_list: | Strd | Orig | ZScore | EScore | Len | Documentation |
|-------------------|------|-------|--------|--------|-----|-------------------------------------|
| gb_est43:AV344569 | - | 29.00 | 133.72 | 110.65 | 205 | AV344569 AV344569 RIKEN full-length |
| gb_est40:AW147227 | - | 29.00 | 132.57 | 128.27 | 235 | AW147227 da16c03.x1 normalized |
| gb_est40:AW147713 | + | 29.00 | 132.18 | 134.78 | 246 | AW147713 da16c03.y1 normalized |
| gb_est44:AW081183 | - | 29.00 | 130.37 | 170.07 | 305 | AW081183 AW081183 Oncorhynchus |
| gb_est41:D36336 | - | 29.00 | 128.97 | 203.48 | 360 | D36336 CELK031FXF Yuji Kohara |
| gb_est24:AI165099 | - | 29.00 | 128.83 | 207.15 | 366 | AI165099 A075p18u Hybrid aspen |
| gb_gss3:B31284 | + | 29.00 | 128.81 | 207.77 | 367 | B31284 HS-1007-A2-H08-MF.ab1 |
| gb_est17:C64475 | - | 29.00 | 128.60 | 213.28 | 376 | C64475 C64475 Yuji Kohara unpub |
| gb_gss5:AQ084632 | + | 29.00 | 128.27 | 222.51 | 391 | AQ084632 HS-2263-B1_E11-T7C.CIT |
| gb_est22:AI044481 | + | 29.00 | 128.21 | 224.35 | 394 | AI044481 UI-R-C1-JW-d-09-O-UI.S |
| gb_est17:AA653822 | - | 29.00 | 128.10 | 227.43 | 399 | AA653822 nt05a11.s1 NCI_CGAP |
| gb_gss8:AA046586 | - | 29.00 | 128.06 | 228.67 | 401 | AA046586 RPI11-3505.TK RPCI-11 |
| gb_est44:AW202244 | - | 29.00 | 127.59 | 242.89 | 424 | AW202244 sf12c06.y1 Gm-cl027 G1 |
| gb_gss15:AQ628040 | + | 29.00 | 127.30 | 252.20 | 439 | AQ628040 nt20064b Medicago trunc |
| gb_gss17:AQ31824 | + | 29.00 | 126.44 | 281.54 | 486 | AQ31824 nbe0045W01r CUGI Rice |
| gb_gss12:AQ379639 | - | 29.00 | 126.13 | 292.84 | 504 | AQ379639 RPI11-135M11.TJ RPCI-1 |
| gb_gss6:AQ866784 | - | 29.00 | 126.05 | 295.98 | 509 | AQ866784 nbe0029E03f CUGI Rice |
| gb_gss13:AQ439640 | - | 29.00 | 125.71 | 309.21 | 530 | AQ439640 HS-5121.A1.F07.T7A RPO |
| gb_est22:AT074354 | - | 29.00 | 125.25 | 327.56 | 559 | AT074354 oz84c12.x1 Soares.sene |
| gb_gss11:AQ303382 | - | 29.00 | 125.12 | 333.27 | 568 | AQ303382 HS-3226.A2.B01.T7 CIT |
| gb_gss11:AQ327017 | + | 29.00 | 124.10 | 379.84 | 641 | AQ327017 nbe0039P10f CUGI Rice |
| gb_gss12:AQ051483 | + | 29.00 | 124.01 | 384.33 | 648 | AQ051483 nbe0002DD03f CUGI Rice |
| gb_est44:AF122505 | + | 29.00 | 123.59 | 405.55 | 681 | AF122505 AF122505 Strongylocent |
| gb_gss11:AQ324361 | + | 29.00 | 123.34 | 419.10 | 702 | AQ324361 mxb0018K05r CUGI Rice |
| gb_est9:AA066472 | + | 28.00 | 135.99 | 82.65 | 88 | AA066472 ml73a10.r1 Stratagene m |
| gb_est10:AI091564 | + | 28.00 | 133.66 | 111.43 | 116 | AI091564 ow59c04.x1 Soares.NSF |
| gb_est10:AI091135 | + | 28.00 | 129.33 | 194.38 | 119 | AI091135 MS46F05.r1 Life Tech m |
| gb_est9:CI7761 | + | 28.00 | 128.54 | 215.06 | 213 | CI7761 CI7761 Human placenta cl |
| gb_est3:R58029 | + | 28.00 | 127.50 | 245.80 | 241 | R58029 F8046 Fetal heart Homo s |
| gb_est11:AA216201 | + | 28.00 | 127.39 | 249.12 | 244 | AA216201 uc06b11.r1 Soares.mous |
| gb_est33:AV097218 | - | 28.00 | 127.36 | 250.22 | 245 | AA216201 uc06b11.r1 Soares.mous |
| gb_est29:AI556167 | + | 28.00 | 126.32 | 285.76 | 277 | AV097218 AV097218 Mus musculus |
| gb_gss9:AA095977 | + | 28.00 | 125.23 | 289.11 | 280 | AI556167 UI-R-C2P-rd-d-05-O-UI. |
| gb_est12:AA305015 | + | 28.00 | 125.96 | 299.18 | 289 | AQ395977 HS-3033.A2.H09.MR CIT |
| gb_est9:AA114469 | + | 28.00 | 125.85 | 303.66 | 293 | AA305015 EST17602 Aorta endothel |
| gb_est13:AA365497 | + | 28.00 | 125.62 | 312.64 | 301 | AA111469 m054c05.r1 Life Tech m |
| gb_est12:AA318337 | + | 28.00 | 125.40 | 321.64 | 309 | AA365497 EST76305 Pineal gland |
| gb_est9:AA098172 | + | 28.00 | 125.16 | 331.79 | 318 | AA38337 EST20441 Retina II Hom |
| gb_est15:AA462897 | + | 28.00 | 124.54 | 358.96 | 342 | AA098172 m86e08.r1 Stratagene |
| gb_est15:AA462897 | + | 28.00 | 124.37 | 366.92 | 349 | AA105858 ml84f07.r1 Stratagene |
| gb_est38:AW073579 | - | 28.00 | 124.32 | 369.19 | 351 | AA462897 v69q02.r1 Soares.mous |
| gb_est38:AW073579 | - | 28.00 | 124.18 | 376.03 | 357 | AW073579 xa40f12.x1 NCI_CGAP_Sa |

gb_est7:W60944 - 28.00 124.02 384.01 364 ! W60944 zd27h01.r1 Soares.fet
 gb_est8:W83178 + 28.00 123.88 390.86 370 ! W83178 mf27g02.r1 Soares.mou
 gb_est36:AI893080 + 28.00 123.83 393.15 372 ! AI893080 ms61g10.y1 Stratage

seq_name: gb_est43:AV344569

seq_documentation_block:

LOCUS AV344569 205 bp mRNA EST 12-NOV-1999
 DEFINITION AV344569 RIKEN full-length enriched, adult male olfactory brain Mus
 musculus cDNA clone 6430545L17 3', mRNA sequence.

ACCESSION AV344569.1 GI:6385628

VERSION AV344569

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 205)

Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
 Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,
 Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I.,
 Kai,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,
 Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
 Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K.,
 Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y.,
 Suzuki,H., Suzuki,H., Takahashi,F., Tateno,M., Tomihara,N.,
 Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A.,
 Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Konno,H., et al.)
 Unpublished (1999)

On May 18, 1998 this sequence version replaced gi:3137512.

JOURNAL

COMMENT
 Contact: Yoshihide Hayashizaki
 Genome Exploration Research Group, Life Science Tsukuba Center,
 The Institute of Physical and Chemical Research (RIKEN), Genomic
 Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: +81-298-36-9013
 Fax: +81-298-36-9098

Email: genome-res@rtc.riken.go.jp,

URL:http://genome.rtc.riken.go.jp/

Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,
 Matsura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and
 Hayashizaki,Y.

Transcriptional sequencing: A method for DNA sequencing using RNA
 polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
 Itoh,M., Katsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
 Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
 Okazaki,Y. and Hayashizaki,Y.

Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)

Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)

Please visit our web site (http://genome.rtc.riken.go.jp) for
 further details.

FEATURES

source

Location/Qualifiers

1. 205

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="6430545L17"

/clone_lib="RIKEN full-length enriched, adult male

olfactory brain"

/sex="male"

/tissue_type="olfactory brain"

/dev_stage="adult"

/lab_host="DH10B"

/note="Site_1: Sall; Site_2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in

RIKEN. Division of Experimental Animal Research in Riken

contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot - 10.0 and subtraction to Rot - 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCGAGTTAATAATATCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: Sall; 3' end: BamHI"

BASE COUNT 46 a 43 c 56 g 60 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-7 x AV344569/rev ..

Align seg 1/1 to reverse of: AV344569 from: 1 to: 205

1 TyrArgLeuAlaIleArg 6
|||||

41 TACAGACTGGCGATTAGG 24

seq_name: gb_est40:AW147227

seq_documentation_block:
LOCUS AW147227 235 bp mRNA EST 30-NOV-1999
DEFINITION dal6c03.x1 normalized Xenopus laevis gastrula Xenopus laevis cDNA
clone XENOPUS_SOURCE_ID:xlnga001n05 3', mRNA sequence.

ACCESSION AW147227
VERSION AW147227.1 GI:6195123

KEYWORDS EST.
SOURCE African clawed frog.

ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae;
Xenopus

REFERENCE 1 (bases 1 to 235)

AUTHORS Johnson,S.L., Blumberg,B., Song,J., Clifton,S., Hillier,L.,
Pape,D., Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y.,
Person,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
Waterston,R. and Willson,R.

TITLE WashU Xenopus EST project, 1999

JOURNAL Unpublished (1999)

COMMENT On Dec 20, 1995 this sequence version replaced gi:1135577.

Other_ESTs: dal6c03.y1

Contact: Stephen L. Johnson/WashU Xenopus EST project, 1999
WashU Xenopus EST project, 1999

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library constructed by Bruce Blumberg

Library normalized by Jihwan Song

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: Xenopus clone distribution information for

this library can be found through Research Genetics, visit their

web page at: <http://www.resgen.com/>

Seq primer: -40UP from Gibco.

FEATURES

source

1..235

/organism="Xenopus laevis"

/db_xref="taxon:8355"

/clone="XENOPUS_SOURCE_ID:xlnga001n05"

/clone_lib="normalized Xenopus laevis gastrula"
/tissue_type="gastrula (stages 10.5, 11.5 mixed)"
/lab_host="Top-10 F"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from 2ug of poly A+ RNA (equal
parts from stage 10.5 and stage 11.5 gastrulae).
EcoRI-XhoI cut cDNA was then ligated into UniZap-XR
(Stratagene) with EcoRI at the 5' end and XhoI at the 3'
end. SS-library phagemids were prepared by mass excision
from the original library and normalized by hybridization
to biotinylated driver (prepared from the same library by
PCR) to Cot-omega of 11. After removal of hybrids and
excess driver by streptavidin sepharose chromatography,
the ss-phagemids were made double stranded and
electroporated into Top-10 F'. Original library
construction by Bruce Blumberg (Cho et al. 1991 Cell 67,
1111-1120). Normalized by Jihwan Song (Song, Cho and
Blumberg, unpublished)."

BASE COUNT 70 a 52 c 45 g 68 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-7 x AW147227/rev ..

Align seg 1/1 to reverse of: AW147227 from: 1 to: 235

1 TyrArgLeuAlaIleArg 6

|||||

205 TACAGATTGGCATACGT 188

seq_name: gb_est40:AW147713

seq_documentation_block:

LOCUS AW147713 246 bp mRNA EST 30-NOV-1999

DEFINITION dal6c03.y1 normalized Xenopus laevis gastrula Xenopus laevis cDNA
clone XENOPUS_SOURCE_ID:xlnga001n05 5', mRNA sequence.

ACCESSION AW147713

VERSION AW147713.1 GI:6195609

KEYWORDS EST.

SOURCE African clawed frog.

ORGANISM Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae;
Xenopus

REFERENCE 1 (bases 1 to 246)

AUTHORS Johnson,S.L., Blumberg,B., Song,J., Clifton,S., Hillier,L.,
Pape,D., Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y.,
Person,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
Waterston,R. and Willson,R.

TITLE WashU Xenopus EST project, 1999

JOURNAL Unpublished (1999)

COMMENT On May 1, 1997 this sequence version replaced gi:2059622.

Other_ESTs: dal6c03.x1

Contact: Stephen L. Johnson/WashU Xenopus EST project, 1999

WashU Xenopus EST project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library constructed by Bruce Blumberg

Library normalized by Jihwan Song

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: Xenopus clone distribution information for

this library can be found through Research Genetics, visit their

web page at: <http://www.resgen.com/>

Seq primer: -40RP from Gibco.

```

FEATURES
  source      Location/Qualifiers
1..246
  /organism="Xenopus laevis"
  /db_xref="taxon:8355"
  /clone="XENOPUS_SOURCE_ID:xlnga001n05"
  /clone_lib="normalized xenopus laevis gastrula"
  /tissue_type="gastrula (stages 10.5, 11.5 mixed)"
  /lab_host="Top-10 F"
  /note="Vector: pluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from Zug of poly A+ RNA (equal
parts from stage 10.5 and stage 11.5 gastrulae).
EcoRI-XhoI cut cDNA was then ligated into Unizap-XR
(Stratagene) with EcoRI at the 5' end and XhoI at the 3'
end. SS-library phagmids were prepared by mass excision
from the original library and normalized by hybridization
to biotinylated driver (prepared from the same library by
PCR) to Cot-omega of 11. After removal of hybrids and
excess driver by streptavidin sepharose chromatography,
the ss-phagmids were made double stranded and
electroporated into Top-10 F'. Original library
constructed by Bruce Blumberg (Cho et al. 1991 Cell 67,
1111-1120). Normalized by Jihwan Song (Song, Cho and
Blumberg, unpublished)."
BASE COUNT      63 a      52 c      59 g      72 t
ORIGIN

alignment_scores:
  Quality: 29.00      Length: 6
  Ratio: 4.833      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-7 x AW147713 ..
|||||
1 TyArgLeuAlaIleArg 6
|||||
63 TACAGATTGGCCATACGT 80

Align seg 1/1 to: AW147713 from: 1 to: 246

seq_name: gb_est44:AU081183

seq_documentation_block:
LOCUS      AU081183      305 bp      mRNA      EST      15-NOV-1999
DEFINITION AU081183 Oncorhynchus mykiss kidney infected by infectious
Hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KN4,
mRNA sequence.
ACCESSION      AU081183
VERSION        AU081183.1 GI:6431531
KEYWORDS       EST.
SOURCE         rainbow trout.
ORGANISM       Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE      1 (bases 1 to 305)
AUTHORS        Sakai, M. and Kono, T.
TITLE          The EST analysis of kidney and gill cells rainbow trout,
Oncorhynchus mykiss, infected with infectious hematopoietic
necrosis virus
JOURNAL        unpublished (1999)
COMMENT        On Jun 5, 1998 this sequence version replaced gi:3189172.
Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: a0b208u@cc.miyazaki-u.ac.jp.
FEATURES
  source      Location/Qualifiers
1..305
  /organism="Oncorhynchus mykiss"
  /db_xref="taxon:8022"
  /clone="KN4"

```

```

/clone_lib="Oncorhynchus mykiss kidney infected by
infectious hematopoietic necrosis virus"
/tissue_type="kidney infected by infectious hematopoietic
necrosis virus"
BASE COUNT      67 a      82 c      87 g      69 t
ORIGIN

alignment_scores:
  Quality: 29.00      Length: 6
  Ratio: 4.833      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-7 x AU081183/rev ..
|||||
1 TyArgLeuAlaIleArg 6
|||||
257 TACCGTTTGGCAATCCG 240

Align seg 1/1 to reverse of: AU081183 from: 1 to: 305

seq_name: gb_est1:D36336

seq_documentation_block:
LOCUS      D36336      360 bp      mRNA      EST      08-AUG-1994
DEFINITION CELK031EXF Yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone YK31f10 5', mRNA sequence.
ACCESSION      D36336
VERSION        D36336.1 GI:528262
KEYWORDS       EST.
SOURCE         Caenorhabditis elegans.
ORGANISM       Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
AUTHORS        Kohara, Y., Mitsuki, H., Nishigaki, A., Motohashi, T., Sugimoto, A. and
Tabara, H.
TITLE          Toward an expression map of the C.elegans genome
JOURNAL        unpublished (1994)
COMMENT        Contact: Yuji Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
  source      Location/Qualifiers
1..360
  /organism="Caenorhabditis elegans"
  /strain="CB1489 him-8(el489)"
  /db_xref="taxon:6239"
  /clone="YK31f10"
  /clone_lib="Yuji Kohara unpublished cDNA"
  /sex="hermaphrodite, male"
  /tissue_type="whole animal"
  /dev_stage="varied"
BASE COUNT      110 a      80 c      87 g      81 t      2 others
ORIGIN

alignment_scores:
  Quality: 29.00      Length: 6
  Ratio: 4.833      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-7 x D36336/rev ..
|||||
1 TyArgLeuAlaIleArg 6
|||||

```

176 TATCGTCTGGCATTCCG 159

seq_name: gb_est24:A1165099

seq_documentation_block:

LOCUS A1165099 366 bp mRNA EST 03-DEC-1998
 DEFINITION A075p18u Hybrid aspen plasmid library Populus tremula x Populus tremuloides cDNA 5', mRNA sequence.

ACCESSION A1165099

VERSION A1165099.1 GI:3856384

KEYWORDS EST.

SOURCE Populus tremula x Populus tremuloides.

ORGANISM Populus tremula x Populus tremuloides

REFERENCE 1 (bases 1 to 366)
 AUTHORS Sterky, F., Regan, S., Karlsson, J., Hertzberg, M., Rohde, A., Holmberg, A., Amini, B., Bhaller, R., Larsson, M., Villarroel, R., Van Montagu, M., Sandberg, G., Olsson, O., Teeri, T., Boerjan, W., Gustafsson, P., Uhlen, M., Sundberg, B. and Lundberg, J.

TITLE Gene discovery in the wood-forming tissues of poplar: Analysis of 5,692 expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (22), 13330-13335 (1998)

MEDLINE 99007314

COMMENT On Jan 17, 1998 this sequence version replaced gi:1900778.
 Contact: Sterky F

Department of Biotechnology

Royal Institute of Technology (KTH)

Teknikringen 34, S-100 44 STOCKHOLM, Sweden

Tel: +46 8 790 8287

Fax: +46 8 24 54 52

Email: fredrik@biochem.kth.se

PCR Primers

FORWARD: AAAGGGGATGCTGCAAGCGG

BACKWARD: GCTTCGGCTGATGTGTGTG

Seq primer: CGTTGAACGACGCCAG

High quality sequence stop: 366.

FEATURES

Source

1..366

/organism="Populus tremula x Populus tremuloides"

/db_xref="taxon:47664"

/clone_lib="Hybrid aspen plasmid library"

/tissue_type="Cambial region"

/dev_stage="1.5 m actively growing tree"

/lab_host="E.coli"

/note="Vector: pBluescript SK; Site_1: SalI; Site_2: NotI; Cambial region tissues, including developing xylem, the meristematic cambial zone and the developing and mature phloem, was harvested from 1.5 m actively growing trees. cDNA was prepared and cloned into lambda gt22a. DNA was isolated and subcloned into pBluescript SK using SalI and NotI restriction enzymes."

BASE COUNT 107 a 63 c 102 g 91 t

ORIGIN

alignment_scores:

Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-7 x A1165099/rev ..

Align seg 1/1 to reverse of: A1165099 from: 1 to: 366

1 TyrArgLeuAlaIleArg 6

188 TATCGCCTGCTATCCGA 171

seq_name: gb_gss3:B31284

seq_documentation_block:

LOCUS B31284 367 bp DNA GSS 17-OCT-1997
 DEFINITION HS-1007-A2-H08-MF.abi CIT Human Genomic Sperm Library C Homo sapiens genomic clone Plate-CT 328 Col-16 Row-O, genomic survey sequence.

ACCESSION B31284

VERSION B31284.1 GI:2530653

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 367)

AUTHORS Mahairas, G.G., Zackrone, K.D., Smith, T., Tipton, S., Schmidt, S., Traicoff, R., Abajian, C., Blanchard, A., West, A. and Hood, L.E.
 TITLE Construction of a Characterized Clone Resource for Genomic Sequencing: Generation and Preliminary Analysis of 20,000 Sequence Tagged Connectors

JOURNAL Unpublished (1997)

COMMENT Contact: Mahairas GG, Zackrone KD, Hood L

University of Washington

Seattle, WA 98195, USA

Tel: (206) 616-8744

Fax: (206) 685-7301

Email: kzackrone@u.washington.edu

Sequence Tagged Connector

Plate: CT 328 row: O column: 16

Class: BAC ends

High quality sequence stop: 367.

Location/Qualifiers

1..367

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="Plate-CT 328 Col-16 Row-O"

/sex="M"

/note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in E-Coli DH10B"

BASE COUNT 87 a 54 c 89 g 137 t

ORIGIN

alignment_scores:

Quality: 29.00 Length: 6

Ratio: 4.833 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-7 x B31284 ..

Align seg 1/1 to: B31284 from: 1 to: 367

1 TyrArgLeuAlaIleArg 6

271 TATCGATTGCCATTCCG 288

seq_name: gb_est17:C64475

seq_documentation_block:

LOCUS C64475 376 bp mRNA EST 22-SEP-1997
 DEFINITION C64475 Yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA clone yk359c8 5', mRNA sequence.

ACCESSION C64475

VERSION C64475.1 GI:2423180

KEYWORDS EST.

SOURCE Caenorhabditis elegans.

ORGANISM Caenorhabditis elegans

REFERENCE 1 (bases 1 to 376)

AUTHORS Kohara, Y., Mochizuki, T., Tabara, H., Watanabe, H., Sugimoto, A., Sano, M., Miyata, A. and Nishigaki, A.

TITLE Expression map of the C.elegans genome

JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1288442.
Contact: Yuji Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
source
1. .376
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/clone_lib="yk359c8"
/clone="yk359c8"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"

BASE COUNT 115 a 83 c 91 g 87 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-7 x C64475/rev ..
Align seg 1/1 to reverse of: C64475 from: 1 to: 376

1 TyrArgLeuAlaIleArg 6
|||||
191 TATCGTCTCGCGATTCGG 174

seq_name: gb_gss5:AQ804632

seq_documentation_block:
LOCUS AQ804632 391 bp DNA GSS 09-AUG-1999
DEFINITION HS_2263-B1_E11-7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-2263 Col-21 Row-J, genomic survey sequence.
ACCESSION AQ804632
VERSION AQ804632.1 GI:5721964
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 391)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 93380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: <http://www.htsc.washington.edu>
Plate: 2263 row: J column: 21
Seq primer: T7
Class: BAC ends
High quality sequence stop: 391.
Location/Qualifiers

source
1. .391
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-2263 Col-21 Row-J"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in E-Coli DH10B"

BASE COUNT 105 a 73 c 65 g 148 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-7 x AQ804632 ..
Align seg 1/1 to: AQ804632 from: 1 to: 391

1 TyrArgLeuAlaIleArg 6
|||||
45 TATAGCTAGCAATTCGA 62

seq_name: gb_est22:AI044481

seq_documentation_block:
LOCUS AI044481 394 bp mRNA EST 11-FEB-1999
DEFINITION UI-R-C1-Jw-d-09-0-UI.s1 UI-R-C1 Rattus norvegicus CDNA clone
UI-R-C1-Jw-d-09-0-UI 3', mRNA sequence.
ACCESSION AI044481
VERSION AI044481.1 GI:3291342
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
REFERENCE 1 (bases 1 to 394)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT On Jan 19, 1998 this sequence version replaced gi:2287562.
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: mscoares@blue.weeg.uiowa.edu
Oligo-dt track not found, Not 1 site shown in beginning of sequence is likely internal to the message. CDNA Library Preparation: M. Fatima Bonaldo, Ph.D. Clone distribution: clones will be available through Research Genetics
Seq primer: M13 Forward

FEATURES
source
1. .394
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C1-Jw-d-09-0-UI"
/clone_lib="UI-R-C1"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-C1 library is a subtracted library derived from the UI-R-C0 library, which is a subtracted library derived from the UI-R-A1 and UI-R-E1 libraries. The UI-R-A1 library consisted of a mixture of individually tagged normalized

libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, and muscle. The UI-R-51 library consisted of a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The substracted library (UI-R-Cl) was constructed as follows: PCR amplified cDNA inserts from UI-R-CO clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-CO library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-Cl library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996).

BASE COUNT 72 a 132 c 108 g 82 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-7 x AI044481 ..

Align seg 1/1 to: AI044481 from: 1 to: 394

1 TyrArgLeuAlaIleArg 6

|||||
297 TACAGACTGGCGATTGCA 314

seq_name: gb_est17:AA653822

seq_documentation_block:
LOCUS AA653822 399 bp mRNA EST 04-NOV-1997
DEFINITION nt05a11.s1 NCI_CGAP_Ov5 Homo sapiens cDNA clone IMAGE:1192220, mRNA sequence.
ACCESSION AA653822
VERSION AA653822.1 GI:2589976
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 399)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1392727.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Andrew Berchuck, M.D., Elise Kohn, M.D., Rodrigo F. Chuacul, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: www-bio.lnl.gov/bbrp/image/image.html

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 398.

Location/Qualifiers

1..399

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1192220"
/clone_lib="NCI_CGAP_Ov5"
/sex="female"
/tissue.type="normal surface epithelium"

/lab_host="DH10B"

/note="Organ: ovary; Vector: pAMP10; mRNA made from normal ovarian epithelium, cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT 65 a 111 c 53 g 170 t
ORIGIN

alignment_scores:

Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-7 x AA653822/rev ..

Align seg 1/1 to reverse of: AA653822 from: 1 to: 399

1 TyrArgLeuAlaIleArg 6

|||||
325 TACAGATTGCGATTAGA 308

seq_name: gb_gss8:AQ046586

seq_documentation_block:

LOCUS AQ046586 401 bp DNA GSS 14-APR-1999
DEFINITION RPC111-3505.TX RPCI-11 Homo sapiens genomic clone RPCI-11-3505, genomic survey sequence.

ACCESSION AQ046586

VERSION AQ046586.1 GI:3315513

KEYWORDS GSS.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 401)

AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.

Use of BAC End Sequences for Sequence-Ready Map Building (1998)

Unpublished (1998)

JOURNAL

COMMENT

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html

Class: BAC ends.

Location/Qualifiers

1..401

source

/organism="Homo sapiens"

/db_xref="GDB:7513396"

/db_xref="taxon:9606"

/clone="RPCI-11-3505"

/clone_lib="RPCI-11"

/sex="Male"

/cell_type="Lymphocytes"

/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
 RPC111 Human Male BAC Library"

111 a 87 c 59 g 144 t

BASE COUNT
 ORIGIN

alignment_scores:
 Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-7 x A0046586/rev ..

Align seg 1/1 to reverse of: A0046586 from: 1 to: 401

1 TyrArgLeuAlaIleArg 6
 267 TATAGATTACCAATTAGA 250

seq_name: gb_est44:AW202244

seq_documentation_block:

LOCUS AW202244 424 bp mRNA 30-NOV-1999
 DEFINITION Gm-c1027-2099 5', mRNA sequence.
 Gm-c1027-2099 5', mRNA sequence.

ACCESSION AW202244

VERSION AW202244.1 GI:6483048

KEYWORDS EST.

SOURCE soybean.

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons: core
 eudicots; Rosidae; eurosids 1; Fabales; Fabaceae; Papilionoideae;
 Glycine.

REFERENCE

AUTHORS

Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,
 Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
 Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
 Bowers,X., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
 Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
 McCann,R., Waterston,R. and Wilson,R.
 Public Soybean EST Project
 Unpublished (1999)

TITLE

JOURNAL

COMMENT

On Jul 7, 1999 this sequence version replaced gi:5866118.

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: Genome Systems, Inc. 4533 World
 Parkway Circle St. Louis, Missouri 63134 For further information
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or
 info@genomesystems.com web site: www.genomesystems.com

High quality sequence stop: 423.

FEATURES

source

Location/Qualifiers

1..424

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1027-2099"

/clone_lib="Gm-c1027"

/tissue_type="cotyledons of 3- and 7-day-old Williams
 seedlings"

/lab_host="DH10B"

/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
 XhoI; This cDNA library was constructed from mRNA isolated
 from cotyledons of 3- and 7-day-old Williams seedlings
 which were propagated on paper towels with distilled
 water. The cotyledons were flash-frozen in liquid
 nitrogen, then lyophilized for 72 hours. Unequal amounts

of mRNA was used for cDNA synthesis. Stratagene's cDNA
 Synthesis Kit (catalog number 200401) was used to
 synthesize the cDNA. First-strand synthesis was
 performed with 5-methyl dCTP, hence the ligated cDNA was
 hemimethylated. A modification of Stratagene's
 first-strand synthesis primer was used. An anchor
 nucleotide (V=A, C, or G) was added to the 3' end of the
 primer [GAGAGAGAGAGAGAGAGACTAGTCTCGAG(T)18] to anchor
 the primer at the 5' end of the poly(A) tract. After
 second-strand synthesis, the cDNA ends were filled in
 with cloned Pfu DNA, ligated to EcoRI adapters and
 subsequently phosphorylated. The XhoI site within the
 first-strand synthesis primer was then restricted by
 digestion with XhoI; all XhoI sites in the cDNA would be
 protected by their hemimethylated status. The cDNA
 constructs were size-fractionated with a 500 bp cutoff,
 using GibcoBRL Life Technologies' cDNA Size Fractionation
 column. The column eluent was then ligated into
 Stratagene's pBluescript(tm) II XR Predigested vector
 (pBluescript II SK(+)) that has been digested with EcoRI
 and XhoI, and phosphorylated by Stratagene). 97% of the
 white and blue colonies appear to contain recombinant
 plasmids with cDNA inserts, based on size (n=30). This
 library was constructed by Dr. Paul Keim and Dr. Virginia
 Coryell."

BASE COUNT 111 a 72 c 112 g 129 t
 ORIGIN

alignment_scores:

Quality: 29.00 Length: 6
 Ratio: 4.833 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-7 x AW202244/rev ..

Align seg 1/1 to reverse of: AW202244 from: 1 to: 424

1 TyrArgLeuAlaIleArg 6
 |||||
 43 TACAGCGCTTCGATCCGT 26

seq_name: gb_gss15:AQ629040

seq_documentation_block:

LOCUS AQ629040 439 bp DNA GSS 27-SEP-1999
 DEFINITION T120064b Medicago truncatula BAC library Medicago truncatula
 genomic clone 03N17, genomic survey sequence.

ACCESSION AQ629040

VERSION AQ629040.1 GI:5091432

KEYWORDS GSS.

SOURCE barrel medic.

ORGANISM

Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicots; Rosidae; eurosids 1; Fabales; Fabaceae; Papilionoideae;
 Medicago.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Mar 23, 1999 this sequence version replaced gi:3323787.

Contact: Cook DR

The Crop Biotechnology Center

Texas A&M University

Department of Plant Pathology and Microbiology, Rm 120 L.F.

Peterson Bldg, College Station, TX 77843-2132, USA

Tel: 409 845 8743

Fax: 409 862 4790

Email: dcooke@pepper.tamu.edu

Other name: 3N177; date: 05/28/99; Submitted to the Database of
 Genome Survey Sequences (GSS) on 06/15/99; More information is

available at 'http://Chryslie.tamu.edu/medicago'; Cultivar: Medicago
truncatula genotype A-17; Note: BAC end.

Seq primer: T7

Class: BAC ends.

FEATURES

source
Location/Qualifiers
1. .439
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="03N17"
/clone_lib="Medicago truncatula BAC library"
/note="vector: pBelOAC11; Site.1: HindIII; Site.2:
HindIII; Nam. Y-W, Penmetza, R.V., Endre, G., Kim, D., and
Cook, D.R. 1999. Construction of a bacterial artificial
chromosome library of Medicago truncatula and
identification of clones containing ethylene response
genes. Theor Appl Genet 98: 638-646."

BASE COUNT 146 a 60 c 81 g 152 t
ORIGIN

alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-7 x AQ629040 ..
Align seg 1/1 to: AQ629040 from: 1 to: 439

1 TyrArgLeuAlaIleArg 6
|||||
206 TATAGATTGGCTATTAGG 223

seq_name: gb_gss7:AQ913824

seq_documentation_block: 486 bp DNA GSS 02-DEC-1999
LOCUS AQ913824
DEFINITION nbeb0045M01r CUGI Rice BAC Library (EcoRI) Oryza sativa genomic
clone nbeb0045M01r, genomic survey sequence.

ACCESSION AQ913824
VERSION AQ913824.1 GI:6510340
KEYWORDS GSS.
SOURCE Oryza sativa.

ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Oryza.
1 (bases 1 to 486)

REFERENCE
AUTHORS Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Rice Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu

Seq primer: GGAACAGCTATGACCATG

Class: BAC ends

High quality sequence start: 66
High quality sequence stop: 403.

Location/Qualifiers

FEATURES

source
1. .486
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="nbeb0045M01r"
/clone_lib="CUGI Rice BAC Library (EcoRI)"
/tissue_type="Leaf"

/lab_host="E. coli DH10B"
/note="vector: pBACIndigo; Site.1: EcoRI; Site.2: EcoRI;
Rice is the most important food crop in the world. Half of
the world population, especially those inhabiting highly
populated areas of the humid tropics and subtropics, rely
on rice as their primary source of carbohydrate.
Monocotyledonous rice is a diploid plant (2n=24) with a
haploid genome equivalent of 431 Mbp (Arumuganathan and
Earle, 1991). The relatively small genome of rice, three
times larger than that of Arabidopsis, makes it suitable
for genomic studies. In order to facilitate positional
cloning, physical mapping and genome sequencing of rice,
we have constructed a BAC library from Oryza sativa,
Nipponbare variety using EcoRI as the cloning enzyme. The
library contains 55,236 clones with an average insert size
of 121 Kb providing approximately 15 haploid genome
equivalents. The deep coverage allows the isolation a
particular sequence with a probability of 99.9 %. Three
high density filters, each containing 18,432 clones
(doubly spotted), represent the whole library for colony
screening and can be requested from the Clemson University
BAC/EST Resource Center (www.genome.clemson.edu)."

BASE COUNT 184 a 78 c 69 g 147 t
ORIGIN
alignment_scores:
Quality: 29.00 Length: 6
Ratio: 4.833 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-7 x AQ913824/rev ..

Align seg 1/1 to reverse of: AQ913824 from: 1 to: 486

1 TyrArgLeuAlaIleArg 6
|||||
258 TACGACTCGCAATTCGA 241

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:29:37 ; Search time 122.56 Seconds
(without alignments)
1.933 Million cell updates/sec

Title: US-08-653-294-8
Perfect score: 49
Sequence: 1 RENLIRALRY 10

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|----------|-----------------------|
| 1 | 49 | 100.0 | 10 | 1 R41208 | Peptide fragment o |
| 2 | 49 | 100.0 | 10 | 1 R3062 | HLA-B2702 CTL modu |
| 3 | 49 | 100.0 | 10 | 1 R95413 | Alphal-helix of HL |
| 4 | 49 | 100.0 | 10 | 1 W07512 | T-cell modulating |
| 5 | 49 | 100.0 | 10 | 1 W47265 | Immunomodulatory p |
| 6 | 49 | 100.0 | 10 | 1 W33784 | Peptide B2702.75-8 |
| 7 | 49 | 100.0 | 10 | 1 R32912 | HLA-B2702 CTL modu |
| 8 | 49 | 100.0 | 15 | 1 W33795 | Peptide B2702.70-8 |
| 9 | 49 | 100.0 | 20 | 1 R92907 | HLA-B2702 CTL modu |
| 10 | 49 | 100.0 | 20 | 1 R92908 | HLA-B2702 CTL modu |
| 11 | 49 | 100.0 | 20 | 1 R95428 | HLA-B2702 84-75-84 |
| 12 | 49 | 100.0 | 20 | 1 W33778 | Immunomodulating d |
| 13 | 49 | 100.0 | 20 | 1 W33791 | Peptide B2702.84-7 |
| 14 | 49 | 100.0 | 25 | 1 R41205 | Peptide fragment o |
| 15 | 49 | 100.0 | 25 | 1 R48286 | Peptide fragment o |
| 16 | 49 | 100.0 | 25 | 1 R83090 | HLA-B2702 CTL modu |
| 17 | 49 | 100.0 | 25 | 1 R83093 | HLA-B2702 CTL modulat |
| 18 | 49 | 100.0 | 25 | 1 R95416 | HLA-B2702.60-84. C |
| 19 | 49 | 100.0 | 25 | 1 R95422 | HLA-B2702.6084. Comps |
| 20 | 49 | 100.0 | 25 | 1 W33794 | Peptide B2702.60-8 |
| 21 | 49 | 100.0 | 184 | 1 Y06801 | Peptide Seq ID No: |
| 22 | 49 | 100.0 | 362 | 1 R03142 | Sequence of HLA-Bw |
| 23 | 49 | 100.0 | 362 | 1 R03144 | Sequence of HLA-B5 |
| 24 | 49 | 100.0 | 362 | 1 R12463 | HLA-Bw53 exon. HLA |
| 25 | 44 | 89.8 | 10 | 1 R30394 | HLA-B2702 CTL modu |
| 26 | 44 | 89.8 | 10 | 1 R30395 | HLA-B2702 CTL modu |
| 27 | 44 | 89.8 | 10 | 1 R30396 | HLA-B2702 CTL modu |
| 28 | 44 | 89.8 | 10 | 1 R95425 | HLA-B2702.75-84(D) |
| 29 | 44 | 89.8 | 10 | 1 R95426 | HLA-B2702.75-84(T) |
| 30 | 44 | 89.8 | 10 | 1 W07513 | T-cell modulating |
| 31 | 44 | 89.8 | 10 | 1 W47267 | Immunomodulatory p |
| 32 | 44 | 89.8 | 10 | 1 W47269 | Immunomodulatory p |
| 33 | 44 | 89.8 | 10 | 1 W33788 | Peptide B2702.75-8 |
| 34 | 44 | 89.8 | 10 | 1 W33787 | Peptide B2702.75-8 |

| | | | | | |
|----|----|------|----|----------|--------------------|
| 35 | 44 | 89.8 | 10 | 1 W33789 | Peptide B2702.75-8 |
| 36 | 44 | 89.8 | 20 | 1 R92909 | HLA-B2702 CTL modu |
| 37 | 44 | 89.8 | 20 | 1 R92910 | HLA-B2702 CTL modu |
| 38 | 44 | 89.8 | 20 | 1 W33792 | Peptide B2702.84-7 |
| 39 | 44 | 89.8 | 20 | 1 W33793 | Peptide B2702.84-7 |
| 40 | 42 | 85.7 | 20 | 1 R95430 | HLA-B2702 84-757/7 |
| 41 | 39 | 79.6 | 10 | 1 R95427 | HLA-B2702.75-84(L) |
| 42 | 39 | 79.6 | 10 | 1 W07514 | T-cell modulating |
| 43 | 39 | 79.6 | 10 | 1 W47271 | Immunomodulatory p |
| 44 | 37 | 75.5 | 10 | 1 W07522 | T-cell modulating |
| 45 | 34 | 69.4 | 10 | 1 R41212 | Peptide fragment o |

ALIGNMENTS

RESULT 1

R41208
ID R41208 standard; peptide; 10 AA.
AC R41208;
DT 15-MAR-1994 (first entry)
DE Peptide fragment of Class I HLA peptide.
KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;
KW parasitic disease; cytotoxic T lymphocyte; modulation.
OS Synthetic.
PN W09317699-A.
PD 16-SEP-1993. U01758.
PF 25-FEB-1993; US-844716.
PR 02-MAR-1992; US-844716.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger CA, Krensky AM;
DR WPI: 93-3031134/38.
PT New peptide(s) based on Class I HLA antigen domains - used for
PT modulating cytotoxic T-lymphocyte activity towards targets
PS Claim 11; Page 54; 61pp; English.
CC The peptide is used to modulate cytotoxic T-lymphocyte (CTL)
CC activity, either by inhibition or stimulation. It can be used
CC for inhibiting CTL toxicity in transplantations, for inducing CTL
CC activity in parasitic diseases and neoplasia and in studies on viral
CC infection. The peptide can also be used for identifying CTLs which
CC bind to it and removing subsets of CTLs from a T-cell composition.
CC This peptide sequence is more commonly found within larger peptide
CC compounds of not more than 30 amino acids in length.
SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RENLIRALRY 10
Db 1 RENLIRALRY 10

RESULT 2

R83062
ID R83062 standard; peptide; 10 AA.
AC R83062;
DT 16-MAY-1996 (first entry)
DE HLA-B2702 CTL modulating peptide (B2702.75-84).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW Class I MHC; HLA-B2702.
OS Synthetic.
PN W09526979-A1.
PD 12-OCT-1995.
PF 05-APR-1995; U04349.
PR 05-APR-1994; US-222851.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM, Farham P;
DR WPI: 95-358582/46.
PT Extension of acceptance period of transplants from MHC unmatched

PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 PT host
 PS Claim 15; Page 9; 80pp; English.
 CC R83061-R83085, R83090-R83096 and R82907-R82913 represent fragments of
 CC class I major histocompatibility complex (MHC) antigens. This sequence
 CC corresponds to residues 75-84 of the alpha-1 domain of the class I MHC
 CC HLA-B*2702. These sequences can be used to extend the period of
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.
 CC The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.00034;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRLALRY 10
 Db | | | | | | | | | |
 1 RENLRLALRY 10

RESULT 3
 R95413
 ID R95413 standard; peptide; 10 AA.
 AC R95413;
 DT 12-NOV-1996 (first entry)
 DE Alpha-helix of HLA-B*2702.
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KW cytolysis; antigen presenting cell.
 OS Synthetic.
 PN W09513288-Al.
 PD 18-MAY-1995.
 PF 10-NOV-1994; U12985.
 PR 10-NOV-1993; US-150493.
 PI (STRD) UNIV LELAND STANFORD JUNIOR.
 PA Clayberger C, Krensky AM;
 DR WPI; 95-194027/25
 PT Compens. comprising lymphoid surface membrane proteins - which may
 PT inhibit cytolytic activity and differentiation of CTLs.
 PS Example; Page 11; 29pp; English.
 CC This sequence represents the alpha-helix of the
 CC human-leucocyte-associated antigen B2702 (HLA-B*2702). This sequence,
 CC epitopes, and palindromes of it (such as R95428) can be used to isolate
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
 CC protein associated with T-cell activation in mammalian T-cells, and is
 CC also immunologically cross reactive with the heat shock protein Hsc70.
 CC p74 is found in a limited number of cell types, but is particularly
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable
 CC cell with an amphoteric detergent, and then passed through an affinity
 CC column containing a covalently bound HLA-B*2702 palindromic peptide.
 CC Compositions comprising the extracellular fragment of p74 combined with
 CC HLA-B*2702.60-84 (see R95416), induces calcium influx, and inhibits
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate
 CC compounds can be screened for their effect on the cytolytic activity of
 CC T-cells, by combining them with the extracellular portion of p74 and
 CC determining the amount of binding between the candidate compound and p74.
 CC Modulation of CTL activity can be inhibited in a cellular composition
 CC containing T-cells and antigen presenting cells (APCs), by adding to the
 CC mix the extracellular portion of p74, in an amount sufficient to compete
 CC with p74 for the binding of the p74 ligand.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.00034;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRLALRY 10
 Db | | | | | | | | | |
 1 RENLRLALRY 10

RESULT 4
 W07512
 ID W07512 standard; peptide; 10 AA.
 AC W07512;
 DT 04-AUG-1997 (first entry)
 DE T-cell modulating peptide #1.
 KW T-cell modulator; autoimmune disease; tissue destruction; alpha-domain;
 KW mammal; major histocompatibility complex; MHC class I; antigen; perforin;
 KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;
 KW rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;
 KW thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;
 KW autologous target cell; cytokine release; T cell activation; therapy.
 OS Synthetic.
 PN W09635443-Al.
 PD 14-NOV-1996.
 PF 05-APR-1996; U04710.
 PR 12-MAY-1995; US-440504.
 PA (SANG-) SANGSTAT MEDICAL CORP.
 PI Buelow R;
 DR WPI; 96-518410/51.
 PT Treatment of auto-immune disease by admin. of peptide(s) corresp. to
 PT major histocompatibility complex antigens - esp. for delaying onset
 PT of clinical symptoms of insulin dependent diabetes by modulating T
 PT cell mediated attack on target cells
 PS Claim 7; Page 20; 24pp; English.
 CC W07512-W07518 represent T-cell modulating peptides that can be used in
 CC the method of the invention. These sequences are based on a portion of
 CC the generic peptide corresponding to residues 70-91 of the alpha-domain
 CC of the major histocompatibility complex (MHC) class I antigen (see
 CC W07510). The method is for affecting the course of an autoimmune disease
 CC involving T-cell mediated destruction of tissue in mammals. These
 CC peptides are used especially to treat insulin-dependent diabetes
 CC mellitus, preferably being administered during the pre-clinical stage to
 CC delay onset of the disease. Other diseases that can be treated are
 CC multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus vulgaris,
 CC Sjogren's disease, thyroid disease, Hashimoto's thyroiditis, myasthenia
 CC gravis, etc. The peptides modulate T-cell mediated attack on autologous
 CC target cells, and may also reduce inflammation, swelling, and release of
 CC cytokines, perforins, granzymes etc. associated with T cell activation.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.00034;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRLALRY 10
 Db | | | | | | | | | |
 1 RENLRLALRY 10

RESULT 5
 W47265
 ID W47265 standard; peptide; 10 AA.
 AC W47265;
 DT 22-MAY-1998 (first entry)
 DE Immunomodulatory peptide.
 KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
 KW transplant rejection; treatment; autoimmune disease.
 OS Homo sapiens.
 OS Synthetic.
 FH Key
 FT Location/Qualifiers
 FT Misc_difference 1..10
 FT /note= "at least one of the amino acids is the
 FT D-isomer
 PN W09744052-Al.
 PD 27-NOV-1997.

PF 23-APR-1997; U06705.
 PR 22-MAY-1996; US-651650.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI; 98-018220/02.
 PT Novel immunomodulatory peptide-type compound - useful for inhibiting
 PT transplant rejection
 PS Claim 10: Page 36; 41pp: English.
 CC The present sequence is an immunomodulatory peptide, which
 CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
 CC in a pharmaceutical composition together with a subtherapeutic dose
 CC of an immunosuppressant, to extend the period of acceptance of a
 CC transplant from a major histocompatibility complex (MHC) unmatched
 CC donor, i.e. to inhibit transplant rejection. It can also be used in
 CC the treatment of autoimmune diseases.
 CC Peptides using the D-form amino acids are more effective
 CC immunomodulators than their diastereomers or enantiomers.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.00034;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRIALRY 10
 DB 1 RENLRIALRY 10

RESULT 6

W33784 ID W33784 standard; peptide; 10 AA.
 AC W33784;
 DT 19-JUN-1998 (first entry)
 DE Peptide B2702.75-84 tested for immunomodulating activity.
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN W09744351-A1.
 PD 27-NOV-1997.
 PF 22-MAY-1997; U08689.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI; 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases

PS Example 1: Page 19; 41pp: English.
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
 CC activity. A peptide-type compound or variant is claimed which has
 CC immunomodulating activity, including the N-terminal acylated and/or
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where
 CC the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha-1 domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.00034;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRIALRY 10
 DB 1 RENLRIALRY 10

RESULT 7

R92912 ID R92912 standard; peptide; 15 AA.
 AC R92912;
 DT 16-MAY-1996 (first entry)
 DE HLA-B2702 CTL modulating peptide (B2702.70-84).
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW class I MHC; HLA-B2702.
 OS Synthetic.
 PN W09526979-A1.
 PD 12-OCT-1995.
 PF 03-APR-1995; U04349.
 PR 05-APR-1994; US-222851.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM, Parham P;
 DR WPI; 95-358582/46.
 PT Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B*5-84 MHC antigen of the recipient
 PT host

PS Example 15: Page 36; 80pp: English.

CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of
 CC class I major histocompatibility complex (MHC) antigens. This sequence
 CC corresponds to residues 70-84 of the alpha-1 domain of the class I MHC
 CC HLA-B2702. These sequences can be used to extend the period of
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.
 CC The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 SQ Sequence 15 AA;

Query Match 100.0%; Score 49; DB 1; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00053;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRIALRY 10
 DB 6 RENLRIALRY 15

RESULT 8

W33795 ID W33795 standard; peptide; 15 AA.
 AC W33795;
 DT 19-JUN-1998 (first entry)
 DE Peptide B2702.70-84 tested for immunomodulating activity.
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN W09744351-A1.
 PD 27-NOV-1997.
 PF 22-MAY-1997; U08689.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI; 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or

PT treating autoimmune diseases
 PS Example 1; Page 19; 41pp; English.
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating activity. A peptide-type compound or variant is claimed which has immunomodulating activity, including the N-terminal acylated and/or C-terminal amidated or esterified forms of up to 60 amino acids, where the peptide-type compound comprises the formula: A-B, where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino acid. The sequence in the brackets may optionally be absent or truncated at any peptide type bond within the brackets. The compounds comprise amino acid sequences related to a Class I HLA-B alpha domain [positions 79-84]. They can be used to inhibit cytotoxic T-lymphocytes (CTL) from undesirably attacking cells in a host or in vitro. They can also be used in combination with antigenic peptides or proteins of interest to activate CTLs. They can also inhibit the proliferation of T cells in response to anti-CD3. The peptide can be used for preventing rejection of transplants or for treating autoimmune diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus. The products can also be used for detection and diagnosis.

Sequence 15 AA;

Query Match 100.0%; Score 49; DB 1; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00053;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLIRALRY 10
 DB 6 RENLIRALRY 15

RESULT 9

R92907
 ID R92907 standard; peptide; 20 AA.
 AC R92907;
 DT 16-MAY-1996 (first entry)
 DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84).
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW class I MHC; HLA-B2702.
 OS Synthetic.
 PN W09526979-A1.
 PD 12-OCT-1995.
 PF 05-APR-1995; U04349.
 PR 05-APR-1994; US-222851.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM, Parham P;
 DR WPI; 95-358582/46.
 PT Extension of acceptance period of transplants from MHC unmatched donor hosts - using Class I B75-84 MHC antigen of the recipient host
 PS Example 15; Page 36; 80pp; English.
 CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of class I major histocompatibility complex (MHC) antigens. This sequence is an inverted dimer of residues 75-84 of the alpha-1 domain of the class I MHC HLA-B2702. These sequences can be used to extend the period of acceptance by a recipient of a transplant from an MHC unmatched donor. The peptides are administered to a patient in conjunction with a subtherapeutic amount of an immunosuppressant. This is administered to the patient for a limited period of time (compared to the lifetime administration for current treatments). The peptides particularly modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs) of the patient.
 Sequence 20 AA;

Query Match 100.0%; Score 49; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.00073;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLIRALRY 10

DB 11 RENLIRALRY 20

RESULT 10

R92908
 ID R92908 standard; peptide; 20 AA.
 AC R92908;
 DT 16-MAY-1996 (first entry)
 DE HLA-B2702 CTL modulating peptide (B2702.84-75(T)/75-84).
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW class I MHC; HLA-B2702.
 OS Synthetic.
 PN W09526979-A1.
 PD 12-OCT-1995.
 PF 05-APR-1995; U04349.
 PR 05-APR-1994; US-222851.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM, Parham P;
 DR WPI; 95-358582/46.
 PT Extension of acceptance period of transplants from MHC unmatched donor hosts - using Class I B75-84 MHC antigen of the recipient host
 PS Example 15; Page 36; 80pp; English.
 CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of class I major histocompatibility complex (MHC) antigens. This sequence is an inverted dimer of residues 75-84 of the alpha-1 domain of the class I MHC HLA-B2702. These sequences can be used to extend the period of acceptance by a recipient of a transplant from an MHC unmatched donor. The peptides are administered to a patient in conjunction with a subtherapeutic amount of an immunosuppressant. This is administered to the patient for a limited period of time (compared to the lifetime administration for current treatments). The peptides particularly modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs) of the patient.
 Sequence 20 AA;

Query Match 100.0%; Score 49; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.00073;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLIRALRY 10
 DB 11 RENLIRALRY 20

RESULT 11

R95428
 ID R95428 standard; peptide; 20 AA.
 AC R95428;
 DT 12-NOV-1996 (first entry)
 DE HLA-B2702 84-75-84 palindromic.
 KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;
 KW T-cell lysate; membrane protein; mammal; heat shock protein; hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KW cytolysis; antigen presenting cell.
 OS Synthetic.
 PN W09513288-A1.
 PD 18-MAY-1995.
 PF 10-NOV-1994; U12985.
 PR 10-NOV-1993; US-150493
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI; 95-194027/25.
 PT Compsns. comprising lymphoid surface membrane proteins - which may inhibit cytolytic activity and differentiation of CTLs.
 PS Example; Page 12; 29pp; English.
 CC R95413, and R95415-R95431 represent palindromes and fragments of human-leucocyte-associated antigens. This sequence represents the HLA-B2702 84-75-84 palindromic. These sequences can be used to isolate the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane

protein associated with T-cell activation in mammalian T-cells, and is also immunologically cross reactive with the heat shock protein Hsc70. p74 is found in a limited number of cell types, but is particularly expressed on B and T cells. p74 can be isolated by lysis of a suitable cell with an amphoteric detergent, and then passed through an affinity column containing a covalently bound HLA-B2702 palindromic peptide. Compositions comprising the extracellular fragment of p74 combined with HLA-B2702 60-84 (see R95416), induces calcium influx, and inhibits cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate compounds can be screened for their effect on the cytolytic activity of T-cells, by combining them with the extracellular portion of p74 and determining the amount of binding between the candidate compound and p74. Modulation of CTL activity can be inhibited in a cellular composition containing T-cells and antigen presenting cells (APCs), by adding to the mix the extracellular portion of p74, in an amount sufficient to compete with p74 for the binding of the p74 ligand.

Sequence 20 AA;

Query Match 100.0%; Score 49; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00073;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RNLRLALRY 10
| | | | | | | |
DB 11 RNLRLALRY 20

RESULT 12

W33778
ID W33778 standard; peptide; 20 AA.
AC W33778;
DT 19-JUN-1998 (first entry)
DE Immunomodulating dimer peptide #1.
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
KW transplanted; autoimmune disease; Class I HLA-B alpha-1 domain;
KW rejection.
OS Synthetic.
OS Homo sapiens.
PN WO9744351-A1.
PD 27-NOV-1997.
PF 22-MAY-1997; U08689.
PR 24-MAY-1996; US-653294.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Beulow R, Clayberger C, Krensky AM;
DR WPI; 98-086530/08.
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B alpha-1 domain, used for preventing rejection of transplants or treating autoimmune diseases
PS Claim 16; Page 35; 41pp; English.

This sequence represents a specifically claimed immunomodulating dimer peptide of the invention. A peptide-type compound or variant is claimed which has immunomodulating activity, including the N-terminal acylated and/or C-terminal amidated or esterified forms of up to 60 amino acids, where the peptide-type compound comprises the formula: A-B, where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino acid. The sequence in the brackets may optionally be absent or truncated at any peptide type bond within the brackets. The compounds comprise amino acids related to a Class I HLA-B alpha domain (positions 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from undesirably attacking cells in a host or in vitro. They can also be used in combination with antigenic peptides or proteins of interest to activate CTLs. They can also inhibit the proliferation of T cells in response to anti-CD3. The peptide can be used for preventing rejection of transplants or for treating autoimmune diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus. The products can also be used for detection and diagnosis.

Sequence 20 AA;

Query Match 100.0%; Score 49; DB 1; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.00073;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RNLRLALRY 10
| | | | | | | |
DB 11 RNLRLALRY 20

RESULT 13

W33791
ID W33791 standard; peptide; 20 AA.
AC W33791;
DT 19-JUN-1998 (first entry)
DE Peptide B2702.84-75/75-84 tested for immunomodulating activity.
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
KW transplanted; autoimmune disease; Class I HLA-B alpha-1 domain;
KW rejection.
OS Synthetic.
OS Homo sapiens.
PN WO9744351-A1.
PD 27-NOV-1997.
PF 22-MAY-1997; U08689.
PR 24-MAY-1996; US-653294.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Beulow R, Clayberger C, Krensky AM;
DR WPI; 98-086530/08.
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B alpha-1 domain, used for preventing rejection of transplants or treating autoimmune diseases
PS Example 1; Page 19; 41pp; English.
CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating activity. A peptide-type compound or variant is claimed which has immunomodulating activity, including the N-terminal acylated and/or C-terminal amidated or esterified forms of up to 60 amino acids, where the peptide-type compound comprises the formula: A-B, where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino acid. The sequence in the brackets may optionally be absent or truncated at any peptide type bond within the brackets. The compounds comprise amino acid sequences related to a Class I HLA-B alpha domain (positions 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from undesirably attacking cells in a host or in vitro. They can also be used in combination with antigenic peptides or proteins of interest to activate CTLs. They can also inhibit the proliferation of T cells in response to anti-CD3. The peptide can be used for preventing rejection of transplants or for treating autoimmune diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus. The products can also be used for detection and diagnosis.

Sequence 20 AA;

Query Match 100.0%; Score 49; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00073;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RNLRLALRY 10
| | | | | | | |
DB 11 RNLRLALRY 20

RESULT 14

R41205
ID R41205 standard; peptide; 25 AA.
AC R41205;
DT 15-MAR-1994 (first entry)
DE Peptide fragment of Class I HLA peptide.
KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;
KW parasitic disease; cytotoxic T lymphocyte; modulation.
OS Synthetic.
PN WO9317699-A.
PD 16-SEP-1993.
PF 25-FEB-1993; U01758.

PR 02-MAR-1992; US-844716.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger CA, Krensky AM;
 DR WPI: 93-303134/38.
 PT New peptide(s) based on Class I HLA antigen domains - used for
 PT modulating cytotoxic T-lymphocyte activity towards targets
 PS Claim 8; Page 53; 61pp; English.
 CC The peptide (or a fragment of at least 10 amino acids, joined at at
 CC least one terminus to a sequence other than that of wild type HLA
 CC antigen) is used to modulate cytotoxic T-lymphocyte (CTL) activity,
 CC either by inhibition or stimulation. It can be used for
 CC inhibiting CTL toxicity in transplantations, for inducing CTL
 CC activity in parasitic diseases and neoplasia and in studies on viral
 CC infection. The peptide can also be used for identifying CTLs which
 CC bind to it and removing subsets of CTLs from a T-cell composition.
 SQ Sequence 25 AA;

Query Match 100.0%; Score 49; DB 1; Length 25;
 Best Local Similarity 100.0%; Pred. No. 0.00094;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRIALRY 10
 Db 16 RENLRIALRY 25
 |||||

RESULT 15
 R48286
 ID R48286 standard; peptide: 25 AA.
 AC R48286;
 DT 15-MAR-1994 (first entry)
 DE Peptide fragment of HLA-B38 antigen.
 KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;
 KW parasitic disease; cytotoxic T lymphocyte; modulation.
 OS Synthetic.
 PN WO9317699-A.
 PD 16-SEP-1993.
 PF 25-FEB-1993; U01758.
 PR 02-MAR-1992; US-844716.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger CA, Krensky AM;
 DR WPI: 93-303134/38.
 PT New peptide(s) based on Class I HLA antigen domains - used for
 PT modulating cytotoxic T-lymphocyte activity towards targets
 PS Example 13; Page 39; 61pp; English.
 CC The peptide is used to modulate cytotoxic T-lymphocyte (CTL)
 CC activity, either by inhibition or stimulation. It can be used for
 CC inhibiting CTL toxicity in transplantations, for inducing CTL
 CC activity in parasitic diseases and neoplasia and in studies on viral
 CC infection. The peptide can also be used for identifying CTLs which
 CC bind to it and removing subsets of CTLs from a T-cell composition.
 CC This peptide is derived from the HLA-B38 antigen and corresponds
 CC to the amino acid positions 60-84 of that antigen.
 SQ Sequence 25 AA;

Query Match 100.0%; Score 49; DB 1; Length 25;
 Best Local Similarity 100.0%; Pred. No. 0.00094;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRIALRY 10
 Db 16 RENLRIALRY 25
 |||||

Search completed: February 8, 2000, 01:29:37
 Job time: 1749 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 11:54:20 ; Search time 117.7 Seconds
(without alignments)
4.008 Million cell updates/sec

Title: US-08-653-294-8
Perfect score: 49
Sequence: 1 RNLRIALRY 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : PIR-62:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 49 | 100.0 | 273 | 2 I38509 | MHC class I histoc |
| 2 | 49 | 100.0 | 274 | 2 I54463 | MHC HLA-B38 chain |
| 3 | 49 | 100.0 | 354 | 2 I59308 | class I histocompa |
| 4 | 49 | 100.0 | 354 | 2 I80168 | class I histocompa |
| 5 | 49 | 100.0 | 354 | 2 I80167 | class I histocompa |
| 6 | 49 | 100.0 | 355 | 2 I80169 | class I histocompa |
| 7 | 49 | 100.0 | 355 | 2 I80171 | class I histocompa |
| 8 | 49 | 100.0 | 359 | 1 HLH012 | MHC class I histoc |
| 9 | 49 | 100.0 | 362 | 1 HLH018 | MHC class I histoc |
| 10 | 49 | 100.0 | 362 | 2 B30345 | MHC class I histoc |
| 11 | 49 | 100.0 | 362 | 2 JH0541 | class I histocompa |
| 12 | 49 | 100.0 | 362 | 2 JH0539 | class I histocompa |
| 13 | 49 | 100.0 | 362 | 2 JH0540 | class I histocompa |
| 14 | 49 | 100.0 | 362 | 2 JH0534 | MHC class I histoc |
| 15 | 49 | 100.0 | 362 | 2 I84486 | transmembrane gly |
| 16 | 49 | 100.0 | 362 | 2 I62045 | gene HLA B-1517 pr |
| 17 | 49 | 100.0 | 362 | 2 I84490 | lymphocyte antigen |
| 18 | 49 | 100.0 | 362 | 2 I37521 | HLA-Bw57.2 antigen |
| 19 | 49 | 100.0 | 362 | 2 A30345 | MHC class I histoc |
| 20 | 49 | 100.0 | 362 | 2 I59633 | MHC HLA-B transmem |
| 21 | 49 | 100.0 | 362 | 2 S24434 | class I histocompa |
| 22 | 49 | 100.0 | 362 | 2 I37120 | MHC class I histoc |
| 23 | 49 | 100.0 | 363 | 2 S07113 | class I histocompa |
| 24 | 49 | 100.0 | 363 | 2 S03537 | class I histocompa |
| 25 | 49 | 100.0 | 364 | 2 D35997 | MHC class I histoc |
| 26 | 49 | 100.0 | 365 | 2 S77963 | MHC class I histoc |
| 27 | 49 | 100.0 | 365 | 2 I54416 | HLA-AW24 protein - |
| 28 | 49 | 100.0 | 365 | 2 I54493 | MHC class I histoc |
| 29 | 44 | 89.8 | 274 | 1 HLH032 | MHC class I histoc |
| 30 | 44 | 89.8 | 355 | 2 I37516 | HLA-B alpha-chain |

```

31      44      89.8      362      2      S25415      class I histocompa
32      44      89.8      362      2      A45850      MHC class I histoc
33      44      89.8      362      2      I61861      MHC HLA-B44.2 chai
34      44      89.8      362      2      I54442      MHC class I histoc
35      44      89.8      364      2      A35997      MHC class I histoc
36      39      79.6      137      2      I80174      class I histocompa
37      39      79.6      359      1      HLH0B4      MHC class I histoc
38      39      79.6      362      2      I54457      MHC class I lympho
39      39      79.6      365      2      JH0537      class I histocompa
40      37      75.5      232      2      I57806      MHC H-2K-kml mRNA
41      37      75.5      362      2      B45876      class I histocompa
42      37      75.5      368      2      A60854      MHC class I histoc
43      37      75.5      368      2      I49712      H-2K-s - mouse
44      37      75.5      368      2      I49713      H-2K-sm1 - mouse
45      37      75.5      369      1      HLMSKK      MHC class I histoc

```

ALIGNMENTS

RESULT 1
I38509

MHC class I histocompatibility antigen - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 23-Jul-1999
C:Accession: I38509
R:Cereb. N.; Choi, J.W.; Riu, K.Z.; Yang, S.Y.
Tissue Antigens 44, 271-273, 1994
A:Title: HLA-B*5105, a newly identified B51 IEF variant.
A:Reference number: I38509; MUID:95176331
A:Accession: I38509
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-273 <RES>
A:Cross-references: EMBL:U06697; NID:g469544; PIDN:AAA92997.1; PID:g469545
C:Genetics:
A:Gene: GDB:HLA-B
A:Cross-references: GDB:120048; OMIM:142830
A:Map position: 6p21.3-6p21.3
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 49; DB 2; Length 273;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RNLRIALRY 10
| | | | | | | | | |
Db 74 RNLRIALRY 83

RESULT 2
I54463

MHC HLA-B38 chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
C:Accession: I54463
R:Mueller, C.A.; Engler-Blum, G.; Gekeler, V.; Steiert, I.; Weiss, E.; Schmidt, H.
Immunogenetics 30, 200-207, 1989
A:Title: Genetic and serological heterogeneity of the supertypic HLA-B locus specific
A:Reference number: I54463; MUID:89379286
A:Accession: I54463
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-274 <RES>
A:Cross-references: GB:M29864; NID:g187674; PIDN:AAA36222.1; PID:g187675
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 49; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RENLIRALRY 10
| | | | |
Db 75 RENLIRALRY 84

RESULT 3

class I histocompatibility antigen - pygmy chimpanzee (fragment)
C:Species: Pan paniscus (pygmy chimpanzee, bonobo)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 23-Jul-1999
C:Accession: I59308
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I59308
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-354 <RES>

A:Cross-references: EMBL:U05575; NID:9454767; PIDN:AAA50178.1; PID:9454768
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 49; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RENLIRALRY 10
| | | | |
Db 91 RENLIRALRY 100

RESULT 4

class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999
C:Accession: I80168
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I80168
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-354 <RES>

A:Cross-references: EMBL:U05579; NID:9454775; PIDN:AAA50182.1; PID:9454776
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 49; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RENLIRALRY 10
| | | | |
Db 91 RENLIRALRY 100

RESULT 5

class I histocompatibility antigen - pygmy chimpanzee (fragment)
C:Species: Pan paniscus (pygmy chimpanzee, bonobo)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999
C:Accession: I80167
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I80167
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-354 <RES>

A:Cross-references: EMBL:U05578; NID:9454773; PIDN:AAA50181.1; PID:9454774
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 49; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RENLIRALRY 10
| | | | |
Db 91 RENLIRALRY 100

RESULT 6

class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999
C:Accession: I80169
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I80169
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-355 <RES>

A:Cross-references: EMBL:U05580; NID:9454777; PIDN:AAA50183.1; PID:9454778
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 49; DB 2; Length 355;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RENLIRALRY 10
| | | | |
Db 91 RENLIRALRY 100

RESULT 7

class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999
C:Accession: I80171
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I80171
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-355 <RES>

A:Cross-references: EMBL:U05582; NID:9454781; PIDN:AAA50185.1; PID:9454782
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 49; DB 2; Length 355;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RENLIRALRY 10
| | | | |
Db 91 RENLIRALRY 100

RESULT 8

MHC class I histocompatibility antigen HLA alpha chain precursor (clone PHLA 12.4)
C:Species: Homo sapiens (man)
C:Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 22-Jun-1999
C:Accession: A02189

R;Malissen, M.; Malissen, B.; Jordan, B.R.
 Proc. Natl. Acad. Sci. U.S.A. 79, 893-897, 1982
 A:Title: Exon/intron organization and complete nucleotide sequence of an HLA gene.
 A:Reference number: A02189; MUID:82151002
 A:Accession: A02189
 A:Molecule type: DNA
 A:Residues: 1-359 <MAL>
 A:Cross-references: GB:J00191; GB:V00526; NID:g187600; PIDN:AAA36218.1; PID:g386873
 C:Comment: The seven exons correspond approximately to the domain structure of this chain
 C:Genetics:
 A:Map position: 6p21.3
 A:Introns: 22/1; 112/1; 204/1; 296/1; 335/1; 346/1
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplanted
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-339/Product: class I histocompatibility antigen HLA alpha chain #status predicted <EXT>
 F:22-304/Domain: extracellular #status predicted <EXT>
 F:22-111/Domain: alpha-1 <EX1>
 F:112-203/Domain: alpha-2 <EX2>
 F:217-282/Domain: immunoglobulin homology <IMM>
 F:305-329/Domain: transmembrane #status predicted <TMN>
 F:335-359/Domain: intracellular #status predicted <INT>
 F:107/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:224-280/Disulfide bonds: #status predicted

Query Match 100.0%; Score 49; DB 1; Length 359;
 Best Local Similarity 100.0%; Pred. No. 0.01;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RENLIRALRY 10
 Db 96 RENLIRALRY 105

RESULT 9
 HLBUB8
 MHC class I histocompatibility antigen HLA-Bw58 alpha chain precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 05-Sep-1997
 C:Accession: A23895
 R;Ways, J.P.; Coppin, H.L.; Parham, P.
 J. Biol. Chem. 260, 11924-11933, 1985
 A:Title: The complete primary structure of HLA-Bw58.
 A:Reference number: A23895; MUID:86008247
 A:Accession: A23895
 A:Molecule type: DNA
 A:Residues: 1-362 <WAY>
 A:Note: the authors translated the codon GCC for residue 349 as Ser
 C:Comment: This protein is a subtype of the HLA-B17 family.
 C:Genetics:
 A:Gene: GDB:HLA-B
 A:Cross-references: GDB:120048; OMIM:142830
 A:Map position: 6p21.3-6p21.3
 A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplanted
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-307/Product: class I histocompatibility antigen HLA-Bw58 alpha chain #status predicted <EXT>
 F:25-307/Domain: extracellular #status predicted <EXT>
 F:25-114/Domain: alpha-1 <EX1>
 F:115-206/Domain: alpha-2 <EX2>
 F:208-331/Domain: immunoglobulin homology <IMM>
 F:308-331/Domain: transmembrane #status predicted <TMN>
 F:332-362/Domain: intracellular #status predicted <INT>
 F:110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 49; DB 1; Length 362;
 Best Local Similarity 100.0%; Pred. No. 0.01;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RENLIRALRY 10

Db 99 RENLIRALRY 108
 RESULT 10
 B30345
 MHC class I histocompatibility antigen HLA-Bw52 precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 16-Feb-1997
 C:Accession: B30345
 R;Hayashi, H.; Ennis, P.D.; Ariga, H.; Salter, R.D.; Parham, P.; Kano, K.; Takiguchi, J.
 Immunol. 142, 306-311, 1989
 A:Title: HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the helical
 A:Reference number: A30345; MUID:89080265
 A:Accession: B30345
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-362 <HAY>
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C:Keywords: transmembrane protein
 F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 49; DB 2; Length 362;
 Best Local Similarity 100.0%; Pred. No. 0.01;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RENLIRALRY 10
 Db 99 RENLIRALRY 108

RESULT 11
 JH0541
 class I histocompatibility antigen Gogo-B0103 heavy chain precursor - lowland gorilla
 C:Species: Gorilla gorilla gorilla (lowland gorilla)
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
 C:Accession: JH0541
 R;Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.
 J. Exp. Med. 174, 1491-1509, 1991
 A:Title: Gorilla class I major histocompatibility complex alleles: comparison to human
 A:Reference number: JH0534; MUID:92078860
 A:Accession: JH0541
 A:Molecule type: DNA
 A:Residues: 1-362 <LAW>
 A:Cross-references: EMBL:X60254; NID:g22869; PIDN:CAA42806.1; PID:g22870
 A:Experimental source: EBV-transformed B cell
 C:Genetics:
 A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C:Keywords: transmembrane protein
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0103 #status
 F:25-114/Domain: alpha-1 <AL1>
 F:115-206/Domain: alpha-2 <AL2>
 F:207-298/Domain: alpha-3 <AL3>
 F:220-285/Domain: immunoglobulin homology <IMM>
 F:229-362/Domain: intracellular #status predicted <INT>

Query Match 100.0%; Score 49; DB 2; Length 362;
 Best Local Similarity 100.0%; Pred. No. 0.01;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RENLIRALRY 10
 Db 99 RENLIRALRY 108

RESULT 12
 JH0539
 class I histocompatibility antigen Gogo-B0101 heavy chain precursor - lowland gorilla
 C:Species: Gorilla gorilla gorilla (lowland gorilla)

C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
 C/Accession: JH0539
 R/Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.
 J. Exp. Med. 174, 1491-1509, 1991
 A/Title: Gorilla class I major histocompatibility complex alleles: comparison to human a
 A/Reference number: JH0534; MUID:92078860
 A/Accession: JH0539
 A/Molecule type: DNA
 A/Residues: 1-362 <LAW>
 A/Cross-references: EMBL:X60255; NID:g22865; PIDN:CAA42807.1; PID:g22866
 A/Experimental source: EBV-transformed B cell
 C/Genetics:
 A/Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1
 C/Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C/Keywords: transmembrane protein
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0101 #status pre
 F:25-114/Domain: alpha-1 <AL1>
 F:115-206/Domain: alpha-2 <AL2>
 F:207-298/Domain: alpha-3 <AL3>
 F:220-285/Domain: immunoglobulin homology <IMM>
 F:299-362/Domain: intracellular #status predicted <INT>

Query Match 100.0% Score 49; DB 2; Length 362;
 Best Local Similarity 100.0% Pred. No. 0.01; Mismatches 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLIRALRY 10
 |||||
 Db 99 RENLIRALRY 108

RESULT 13

JH0540
 Class I histocompatibility antigen Gogo-B0102 heavy chain precursor - lowland gorilla
 C/Species: Gorilla gorilla gorilla (lowland gorilla)
 C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
 C/Accession: JH0540
 R/Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.
 J. Exp. Med. 174, 1491-1509, 1991
 A/Title: Gorilla class I major histocompatibility complex alleles: comparison to human a
 A/Reference number: JH0534; MUID:92078860
 A/Accession: JH0540
 A/Molecule type: DNA
 A/Residues: 1-362 <LAW>
 A/Cross-references: EMBL:X60693; NID:g22867; PIDN:CAA43101.1; PID:g22868
 A/Experimental source: EBV-transformed B cell
 C/Genetics:
 A/Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1
 C/Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C/Keywords: transmembrane protein
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0102 #status pre
 F:25-114/Domain: alpha-1 <AL1>
 F:115-206/Domain: alpha-2 <AL2>
 F:207-298/Domain: alpha-3 <AL3>
 F:220-285/Domain: immunoglobulin homology <IMM>
 F:299-362/Domain: intracellular #status predicted <INT>

Query Match 100.0% Score 49; DB 2; Length 362;
 Best Local Similarity 100.0% Pred. No. 0.01; Mismatches 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLIRALRY 10
 |||||
 Db 99 RENLIRALRY 108

RESULT 14

A45834
 MHC class I histocompatibility antigen HLA-B53 alpha chain precursor - human

C/Species: Homo sapiens (man)
 C/Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 23-Jul-1999
 C/Accession: A45834
 R/Hayashi, H.; Ooba, T.; Nakayama, S.; Sekimata, M.; Kano, K.; Takiguchi, M.
 Immunogenetics 32, 195-199, 1990
 A/Title: Allospecificities between HLA-Bw53 and HLA-B35 are generated by substitution
 A/Reference number: A45834; MUID:91033941
 A/Accession: A45834
 A/Molecule type: DNA
 A/Residues: 1-362 <HAY>
 A/Cross-references: GB:M58636; NID:g187756; PIDN:AAA36228.1; PID:g187757; GB:M33574
 A/Note: this allele is designated B*5301
 C/Genetics:
 A/Gene: GDB:HLA-B
 A/Cross-references: GDB:120048; OMIM:142830
 A/Map position: 6p21.3-6p21.3
 C/Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C/Keywords: glycoprotein; heterodimer; transmembrane protein
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:220-285/Domain: immunoglobulin homology <IMM>
 F:110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0% Score 49; DB 2; Length 362;
 Best Local Similarity 100.0% Pred. No. 0.01; Mismatches 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLIRALRY 10
 |||||
 Db 99 RENLIRALRY 108

RESULT 15

I8486
 transmembrane glycoprotein - human
 C/Species: Homo sapiens (man)
 C/Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 23-Jul-1999
 C/Accession: I8486
 R/Hildebrand, W.H.; Domene, J.D.; Shen, S.Y.; Lau, M.; Terasaki, P.I.; Bunce, M.; Mar
 Tissue Antigens 43, 209-218, 1994
 A/Title: HLA-B15: a widespread and diverse family of HLA-B alleles.
 A/Reference number: I38421; MUID:94367483
 A/Accession: I8486
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-362 <RES>
 A/Cross-references: GB:LI5005; NID:g493154; PIDN:AAA56832.1; PID:g493155
 C/Genetics:
 A/Gene: HLA-B*1513
 C/Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C/Keywords: glycoprotein

Query Match 100.0% Score 49; DB 2; Length 362;
 Best Local Similarity 100.0% Pred. No. 0.01; Mismatches 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLIRALRY 10
 |||||
 Db 99 RENLIRALRY 108

Search completed: February 7, 2000, 11:54:20
 Job time: 24330 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 00:59:49 ; Search time 63.71 Seconds
(without alignments)
4.688 Million cell updates/sec

Title: US-08-653-294-8
Perfect score: 49
Sequence: 1 RENLIALRY 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_38.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------------|--------------------|
| 1 | 49 | 100.0 | 359 | 1B01_PANTR | P13750 pan troglod |
| 2 | 49 | 100.0 | 362 | 1B01_GORGO | P30379 gorilla gor |
| 3 | 49 | 100.0 | 362 | 1B02_GORGO | P30380 gorilla gor |
| 4 | 49 | 100.0 | 362 | 1B03_GORGO | P30381 gorilla gor |
| 5 | 49 | 100.0 | 362 | 1B15_HUMAN | P10317 homo sapien |
| 6 | 49 | 100.0 | 362 | 1B47_HUMAN | P30487 homo sapien |
| 7 | 49 | 100.0 | 362 | 1B49_HUMAN | P18484 homo sapien |
| 8 | 49 | 100.0 | 362 | 1B52_HUMAN | P30489 homo sapien |
| 9 | 49 | 100.0 | 362 | 1B53_HUMAN | P30490 homo sapien |
| 10 | 49 | 100.0 | 362 | 1B54_HUMAN | P30491 homo sapien |
| 11 | 49 | 100.0 | 362 | 1B60_HUMAN | P18465 homo sapien |
| 12 | 49 | 100.0 | 362 | 1B61_HUMAN | P30497 homo sapien |
| 13 | 49 | 100.0 | 362 | 1B62_HUMAN | P10319 homo sapien |
| 14 | 49 | 100.0 | 362 | 1B62_HUMAN | P10319 homo sapien |
| 15 | 49 | 100.0 | 365 | 1A23_HUMAN | P30447 homo sapien |
| 16 | 49 | 100.0 | 365 | 1A24_HUMAN | P05534 homo sapien |
| 17 | 44 | 89.8 | 362 | 1B05_HUMAN | P30461 homo sapien |
| 18 | 44 | 89.8 | 362 | 1B41_HUMAN | P30481 homo sapien |
| 19 | 44 | 89.8 | 362 | 1B42_HUMAN | P30482 homo sapien |
| 20 | 44 | 89.8 | 365 | 1A25_HUMAN | P18462 homo sapien |
| 21 | 44 | 89.8 | 365 | 1A32_HUMAN | P10314 homo sapien |
| 22 | 39 | 79.6 | 359 | 1B40_HUMAN | P10320 homo sapien |
| 23 | 39 | 79.6 | 365 | 1A04_GORGO | P30378 gorilla gor |
| 24 | 37 | 75.5 | 369 | 1H1K_MOUSE | P04223 mus musculu |
| 25 | 34 | 69.4 | 299 | 1VF19_HSV6U | P52348 herpes simp |
| 26 | 34 | 69.4 | 338 | 1B20_HUMAN | P30467 homo sapien |
| 27 | 34 | 69.4 | 361 | 1B14_HUMAN | P03989 homo sapien |
| 28 | 34 | 69.4 | 362 | 1B16_HUMAN | P19373 homo sapien |
| 29 | 34 | 69.4 | 362 | 1B18_HUMAN | P10318 homo sapien |
| 30 | 34 | 69.4 | 362 | 1B19_HUMAN | Q08136 homo sapien |
| 31 | 34 | 69.4 | 362 | 1B29_HUMAN | P18463 homo sapien |
| 32 | 34 | 69.4 | 362 | 1B45_HUMAN | P30485 homo sapien |
| 33 | 33 | 67.3 | 533 | 1MASV_ECOLI | P08997 escherichia |
| 34 | 32 | 65.3 | 298 | 1HALY_MOUSE | P01895 mus musculu |

RESULT 1

| ID | 1B01_PANTR | STANDARD; | PRT; | 359 AA. |
|----|---|-----------|------|---|
| AC | P13750; | | | |
| DT | 01-JAN-1990 (Rel. 13, Created) | | | |
| DT | 01-JAN-1990 (Rel. 13, Last sequence update) | | | |
| DT | 01-APR-1993 (Rel. 23, Last annotation update) | | | |
| DE | CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-1 ALPHA CHAIN PRECURSOR (FRAGMENT). | | | |
| OS | Pan troglodytes (Chimpanzee). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; | | | |
| OC | Eutheria; Primates; Catarrhini; Homnidae; Pan. | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE; 89030641. | | | |
| RA | MAYER W.E., JONKER M., KLEIN D., IVANYI P., VAN SEVENTER G., KLEIN J. | | | |
| RT | "Nucleotide sequences of chimpanzee MHC class I alleles: evidence for trans-species mode of evolution." | | | |
| RL | EMBO J. 7:2765-2774(1988). | | | |
| RN | [2] | | | |
| RP | REVISIONS. | | | |
| RA | MAYER W. | | | |
| RL | Submitted (FEB-1989) to the EMBL/GenBank/DBJ databases. | | | |
| CC | -I- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE IMMUNE SYSTEM. | | | |
| CC | -I- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-MICROGLOBULIN). | | | |
| CC | THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | | | |
| CC | EMBL; X13115; CAA31507.1; . | | | |
| DR | PIR; S03537; S03537. | | | |
| DR | HSSP; P03989; IHS. | | | |
| DR | PROSITE; PS00290; IG_MHC; 1. | | | |
| DR | PFAM; PF00047; Ig; 1. | | | |
| DR | PFAM; PF00129; MHC_I; 1. | | | |
| KW | MHC I; transmembrane; Glycoprotein; Signal. | | | |
| FT | NON_TER | 1 | 1 | |
| FT | SIGNAL | <1 | 20 | |
| FT | CHAIN | 21 | 359 | |
| FT | DOMAIN | 21 | 110 | CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-1 ALPHA CHAIN. |
| FT | DOMAIN | 111 | 202 | EXTRACELLULAR ALPHA-1. |
| FT | DOMAIN | 203 | 294 | EXTRACELLULAR ALPHA-2. |
| FT | DOMAIN | 295 | 305 | EXTRACELLULAR ALPHA-3. |
| FT | TRANSNEM | 306 | 329 | CONNECTING PEPTIDE. |
| FT | DOMAIN | 330 | 359 | CYTOPLASMIC TAIL. |
| FT | DISULFID | 121 | 184 | BY SIMILARITY. |
| FT | DISULFID | 223 | 279 | BY SIMILARITY. |
| FT | CARBOHYD | 106 | 106 | BY SIMILARITY. |

35 32 65.3 360 1 YD95_MYCTU
36 32 65.3 361 1 HALA_RABIT
37 32 65.3 361 1 HALB_RABIT
38 32 65.3 363 1 LACK_AGRD
39 32 65.3 365 1 HAL2_MOUSE
40 32 65.3 616 1 YJGL_ECOLI
41 32 65.3 843 1 VAV_RAT
42 32 65.3 845 1 VAV_MOUSE
43 31 63.3 228 1 YMEB_MTEEX
44 31 63.3 244 1 YEHT_ECOLI
45 31 63.3 268 1 YRY3_CAEEL

ALIGNMENTS

P71663 mycobacteri
P01894 oryctolagus
P06140 oryctolagus
Q01937 agrobacteri
P01900 mus musculu
P39336 escherichia
P54100 rattus norv
P27870 mus musculu
Q49116 methylobact
P33356 escherichia
Q10007 caenorhabdi

SQ SEQUENCE 359 AA; 40173 MW; 5395FFC9 CRC32;

Query Match 100.0%; Score 49; DB 1; Length 359;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENRLALRY 10
| | | | |
DB 95 RENRLALRY 104

RESULT 2
ID 1B01_GORGO STANDARD; PRT; 362 AA.
AC P30379;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0101 ALPHA CHAIN PRECURSOR.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92078860.
RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;
RT "Gorilla class I major histocompatibility complex alleles: comparison
to human and chimpanzee class I.";
RL J. Exp. Med. 174:1491-1509(1991).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60255; CAA42807.1; -
DR PIR; JH0539; JH0539.
DR HSSP; P03989; ILSA.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362
FT DOMAIN 25 114
FT DOMAIN 115 206
FT DOMAIN 207 298
FT DOMAIN 299 308
FT TRANSMEM 309 332
FT DOMAIN 333 362
FT DISULFID 125 188
FT DISULFID 227 283
FT CARBOHYD 110 110
SQ SEQUENCE 362 AA; 40170 MW; 2E33E2B8 CRC32;

Query Match 100.0%; Score 49; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENRLALRY 10
| | | | |
DB 99 RENRLALRY 108

RESULT 3
ID 1B02_GORGO STANDARD; PRT; 362 AA.
AC P30380;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0102 ALPHA CHAIN PRECURSOR.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92078860.
RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;
RT "Gorilla class I major histocompatibility complex alleles: comparison
to human and chimpanzee class I.";
RL J. Exp. Med. 174:1491-1509(1991).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60693; CAA43101.1; -
DR PIR; JH0540; JH0540.
DR HSSP; P03989; ILSA.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362
FT DOMAIN 25 114
FT DOMAIN 115 206
FT DOMAIN 207 298
FT DOMAIN 299 308
FT TRANSMEM 309 332
FT DOMAIN 333 362
FT DISULFID 125 188
FT DISULFID 227 283
FT CARBOHYD 110 110
SQ SEQUENCE 362 AA; 40204 MW; 3CF119AD CRC32;

Query Match 100.0%; Score 49; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENRLALRY 10
| | | | |
DB 99 RENRLALRY 108

RESULT 4
ID 1B03_GORGO STANDARD; PRT; 362 AA.
AC P30381;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0103 ALPHA CHAIN PRECURSOR.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92078860.
 RA LAWLER D.A., WARREN E., TAYLOR P., PARHAM P.;
 RT "Gorilla class I major histocompatibility complex alleles: comparison
 to human and chimpanzee class I.";
 RL J. Exp. Med. 174:1491-1509(1991).
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 THE IMMUNE SYSTEM.
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 MICROGLOBULIN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 DR EMBL: X60254; CAA42806.1; -
 DR PIR: JH0541; JH0541.
 DR HSP: P03989; IHSA.
 DR PFAM: PF00047; ig: 1.
 DR PROSITE: PS00290; IG_MHC; 1.
 DR PFAM: PF00129; MHC_1; 1.
 KW MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT DOMAIN 309 332
 FT DOMAIN 333 362
 FT DISULFID 125 188
 FT DISULFID 227 283
 FT CARBOHYD 110 110
 SQ SEQUENCE 362 AA; 40248 MW; FEA6A941 CRC32;
 Query Match 100.0%; Score 49; DB 1; Length 362;
 Best Local Similarity 100.0%; Pred. No. 0.0037;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 RENLIRALRY 10
 Db 99 RENLIRALRY 108
 RESULT 5
 ID 1B15_HUMAN STANDARD; PRT; 362 AA.
 AC P10317;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-27 B*2702 ALPHA CHAIN
 DE PRECURSOR (B-27K) (B27.2).
 GN HLA-B OR HLAB
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
 CC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 86220133.
 RA SEEMANN G.H.A., REIN R.S., BROWN C.S., PLOEGH H.L.;
 RT "Gene conversion-like mechanisms may generate polymorphism in human
 class I genes.";
 RL EMBO J. 5:547-552(1986).
 RN [2]

RP SEQUENCE FROM N.A.
 RA PARHAM P., ARNETT K.L., ADAMS E.J.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 86-107 AND 171-181.
 RX MEDLINE: 86042671.
 RA VEGA M.A., EZQUERRA A., ROJO S., APARICIO P., BRAGADO R.,
 RA LOPEZ DE CASTRO J.A.;
 RT "Structural analysis of an HLA-B27 functional variant: identification
 of residues that contribute to the specificity of recognition by
 cytolytic T lymphocytes.";
 RT Proc. Natl. Acad. Sci. U.S.A. 82:7394-7398(1985).
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 THE IMMUNE SYSTEM.
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 MICROGLOBULIN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 DR EMBL: X03664; CAA27301.1; -
 DR EMBL: X03667; CAA27301.1; JOINED.
 DR EMBL: L38504; AAG69724.1; -
 DR PIR: B25092; HLHOBK.
 DR HSP: P03989; IHSA.
 DR MIM: 142830; -
 DR PROSITE: PS00290; IG_MHC; 1.
 DR PFAM: PF00047; ig: 1.
 DR PFAM: PF00129; MHC_1; 1.
 KW MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT DOMAIN 309 332
 FT DOMAIN 333 362
 FT CARBOHYD 110 110
 FT DISULFID 125 188
 FT DISULFID 227 283
 SQ SEQUENCE 362 AA; 40397 MW; 9798F0BB CRC32;
 Query Match 100.0%; Score 49; DB 1; Length 362;
 Best Local Similarity 100.0%; Pred. No. 0.0037;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 RENLIRALRY 10
 Db 99 RENLIRALRY 108
 RESULT 6
 ID 1B47_HUMAN STANDARD; PRT; 362 AA.
 AC P30487;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-49(B-21) B*4901 ALPHA CHAIN
 DE PRECURSOR.
 GN HLA-B OR HLAB
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
 CC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]


```

DE DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-52(B-5) B*5201 ALPHA CHAIN
DE DE PRECURSOR.
CN CN HLA-B OR HLAB.
OS OS Homo sapiens (Human).
OC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
RN RN Eutheria; Primates; Catarrhini; Homnidae; Homo.
[1]
RP RP SEQUENCE FROM N.A.
RX RX MEDLINE; 89080265.
RA RA HAYASHI H., ENNIS P.D., ARIGA H., SALTER R.D., PARHAM P., KANO K.,
RA TAKIGUCHI M.;
RT RT "HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the
RT helical region of the alpha 1 domain.";
RL J. Immunol. 142:306-311(1989).
CC CC -I- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC CC -I- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
CC CC EMBL; M22799; AAA59645.1; ALT_SEQ.
DR DR EMBL; M22793; AAA59645.1; JOINED.
DR DR EMBL; M22794; AAA59645.1; JOINED.
DR DR EMBL; M22795; AAA59645.1; JOINED.
DR DR EMBL; M22796; AAA59645.1; JOINED.
DR DR EMBL; M22797; AAA59645.1; JOINED.
DR DR EMBL; M22798; AAA59645.1; JOINED.
DR DR PIR; B30345; B30345.
DR DR HSP; P30491; IAI1.
DR DR MIM; 142830.
DR DR PROSITE; PS00290; IG_MHC; 1.
DR DR PFAM; PF00047; 19; 1.
DR DR PFAM; PF00129; MHC_I; 1.
DR DR MHC I; Transmembrane; Glycoprotein; Signal.
FT FT SIGNAL 1 24
FT FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT FT B-51(B-5) B*5104 ALPHA CHAIN.
FT FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
FT FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
FT FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
FT FT DOMAIN 299 308 CONNECTING PEPTIDE.
FT FT TRANSMEM 309 332
FT FT DOMAIN 333 362 CYTOPLASMIC TAIL.
FT FT CARBOHYD 110 110 BY SIMILARITY.
FT FT DISULFID 125 188 BY SIMILARITY.
FT FT DISULFID 227 283 BY SIMILARITY.
SQ SEQUENCE 362 AA; 40560 MW; F22F08AB CRC32;

Query Match 100.0%; Score 49; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLIRALRY 10
Db 99 RENLIRALRY 108
.

RESULT 9
IB53_HUMAN
ID IB53_HUMAN STANDARD; PRT; 362 AA.
AC P30490;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)

DE DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-51(B-5) B*5104 ALPHA CHAIN
DE DE PRECURSOR.
CN CN HLA-B OR HLAB.
OS OS Homo sapiens (Human).
OC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
RN RN Eutheria; Primates; Catarrhini; Homnidae; Homo.
[1]
RP RP SEQUENCE FROM N.A.
RX RX MEDLINE; 92269955.
RA RA BELICH M.P., MADRIGAL J.A., HILDEBRAND W.H., ZEMMOUR J.,
RA WILLIAMS R.C., LUZ R., PETZL-ERLER M.L., PARHAM P.;
RT RT "Unusual HLA-B alleles in two tribes of Brazilian Indians.";
RL Nature 357:326-329(1992).
CC CC -I- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC CC -I- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
CC CC EMBL; Z15143; CAA78849.1; -.
DR DR HSP; P30491; IAI1.
DR DR MIM; 142830.
DR DR PROSITE; PS00290; IG_MHC; 1.
DR DR PFAM; PF00047; 19; 1.
DR DR PFAM; PF00129; MHC_I; 1.
DR DR MHC I; Transmembrane; Glycoprotein; Signal.
FT FT SIGNAL 1 24
FT FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT FT B-51(B-5) B*5104 ALPHA CHAIN.
FT FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
FT FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
FT FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
FT FT DOMAIN 299 308 CONNECTING PEPTIDE.
FT FT TRANSMEM 309 332
FT FT DOMAIN 333 362 CYTOPLASMIC TAIL.
FT FT CARBOHYD 110 110 BY SIMILARITY.
FT FT DISULFID 125 188 BY SIMILARITY.
FT FT DISULFID 227 283 BY SIMILARITY.
SQ SEQUENCE 362 AA; 40560 MW; F22F08AB CRC32;

Query Match 100.0%; Score 49; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLIRALRY 10
Db 99 RENLIRALRY 108
.

RESULT 10
IB54_HUMAN
ID IB54_HUMAN STANDARD; PRT; 362 AA.
AC P30491;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)

```

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-53 B*5301 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91033941.
RA HAYASHI H., OBOA T., NAKAYAMA S., SEKIMATA M., KANO K.,
RA TAKIGUCHI M.;
RT "Allospecificities between HLA-B*53 and HLA-B*35 are generated by
RT substitution of the residues associated with HLA-B*4/B*6 public
RT epitopes";
RL Immunogenetics 32:195-199(1990).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 25-302.
RX MEDLINE: 96209672.
RA SMITH K.J., REID S.W., HARLOS K., MCMICHAEL A.J., STUART D.I.,
RA BELL J.I., JONES E.Y.;
RT "Bound water structure and polymorphic amino acids act together to
RT allow the binding of different peptides to MHC class I HLA-B*53";
RL Immunity 4:215-228(1996).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M58636; AAA36228.1; -
DR PIR: A45834; A45834.
DR PDB: 1A1M; 08-APR-98.
DR PDB: 1A1O; 08-APR-98.
DR MIM: 142830; -
DR PROSITE: PS00290; IG_MHC; 1.
DR PFAM: PF00047; Ig; 1.
DR PFAM: PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT BW-53 B*5301 ALPHA CHAIN.
FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
FT DOMAIN 299 308 CONNECTING PEPTIDE.
FT TRANSMEM 309 332
FT DOMAIN 333 362 CYTOPLASMIC TAIL.
FT CARBOHYD 110 110 BY SIMILARITY.
FT DISULFID 125 188
FT DISULFID 227 283
SQ SEQUENCE 362 AA; 40495 MW; 2BDC746E CRC32;
Query Match 100.0%; Score 49; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RENLRLRY 10
Db 99 RENLRLRY 108
RESULT 11
ID 1B60_HUMAN STANDARD; PRT; 362 AA.
AC P18465;

DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B*57(B-17) B*5701 ALPHA
DE CHAIN PRECURSOR (BW57.1).
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 90207291.
RA ENNIS P.D., ZEMOUR J., SALTER R.D., PARRAM P.;
RT "Rapid cloning of HLA-A,B cDNA by using the polymerase chain
RT reaction: frequency and nature of errors produced in amplification";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91067476.
RA ISAMAT M., GIRDLESTONE J., MILSTEIN C.;
RT "Nucleotide sequence of an HLA-B*57 gene";
RL Nucleic Acids Res. 18:6702-6702(1990).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M32318; AAA36231.1; -
DR EMBL: X55711; CRA39244.1; -
DR PIR: S12622; S12622.
DR PIR: D35997; D35997.
DR HSP: P30491; 1A1M.
DR MIM: 142830; -
DR PROSITE: PS00290; IG_MHC; 1.
DR PFAM: PF00047; Ig; 1.
DR PFAM: PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT B*57(B-17) B*5701 ALPHA CHAIN.
FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
FT DOMAIN 299 308 CONNECTING PEPTIDE.
FT TRANSMEM 309 332
FT DOMAIN 333 362 CYTOPLASMIC TAIL.
FT CARBOHYD 110 110 BY SIMILARITY.
FT DISULFID 125 188 BY SIMILARITY.
FT DISULFID 227 283 BY SIMILARITY.
SQ SEQUENCE 362 AA; 40224 MW; D91DF8DD CRC32;
Query Match 100.0%; Score 49; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RENLRLRY 10
Db 99 RENLRLRY 108
RESULT 12
ID 1B61_HUMAN STANDARD; PRT; 362 AA.
AC P30497;

DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-57(B-17) B*5702 ALPHA CHAIN
 DE PRECURSOR (BW57.2).
 GN HLA-B OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 93056508.
 RA MADRICAL J.A., BELICH M.P., HILDEBRAND W.H., BENJAMIN R.J.,
 RA LITTLE A.M., ZEMOUR J., ENNIS P.D., WARD F.E., PETZL-ERLER M.L.,
 RA MARTEL R.W., DU TOIT E.D., PARHAM P.;
 RT "Distinctive HLA-A,B antigens of black populations formed by
 RT interallelic conversion.";
 RL J. Immunol. 149:3411-3415(1992).
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X61707; CAA43876.1; -;
 DR PIR; S16774; S16774.
 DR HSP; P30491; IALM.
 DR MIM; 142830; -;
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; ig; 1.
 DR PFAM; PF00129; MHC_I; 1.
 DR MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT TRANSMEM 309 332
 FT DOMAIN 333 362
 FT CARBOHYD 110 110
 FT DISULFID 125 188
 FT DISULFID 227 283
 SQ SEQUENCE 362 AA; 40342 MW; 628C2156 CRC32;
 Query Match 100.0%; Score 49; DB 1; Length 362;
 Best Local Similarity 100.0%; Pred. No. 0.0037;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RENLIRALRY 10
 DB 99 RENLIRALRY 108
 RESULT 13
 ID 1B62_HUMAN STANDARD; PRT; 362 AA.
 AC P10319;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-58(B-17) B*5801 ALPHA
 DE CHAIN PRECURSOR.
 GN HLA-B OR HLAB.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86008247.
 RA WAYS J.P., COPPIN H.L., PARHAM P.;
 RT "The complete primary structure of HLA-Bw58.";
 RL J. Biol. Chem. 260:11924-11933(1985).
 CC [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86008247.
 RA INOUE T., OGAWA A.;
 RA Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M11799; AAS59628.1; -;
 DR EMBL; AB008102; BAA22916.1; -;
 DR PIR; A23895; HLHUB8.
 DR HSP; P30491; IALM.
 DR MIM; 142830; -;
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; ig; 1.
 DR PFAM; PF00129; MHC_I; 1.
 DR MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT TRANSMEM 309 332
 FT DOMAIN 333 362
 FT CARBOHYD 110 110
 FT DISULFID 125 188
 FT DISULFID 227 283
 SQ SEQUENCE 362 AA; 40337 MW; 3E5E7534 CRC32;
 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
 BW-58(B-17) B*5801 ALPHA CHAIN.
 EXTRACELLULAR ALPHA-1.
 EXTRACELLULAR ALPHA-2.
 EXTRACELLULAR ALPHA-3.
 CONNECTING PEPTIDE.
 CYTOPLASMIC TAIL.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 Query Match 100.0%; Score 49; DB 1; Length 362;
 Best Local Similarity 100.0%; Pred. No. 0.0037;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RENLIRALRY 10
 DB 99 RENLIRALRY 108
 RESULT 14
 ID HLAH_HUMAN STANDARD; PRT; 362 AA.
 AC P01893;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-DEC-1999 (Rel. 39, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, ALPHA CHAIN H PRECURSOR
 DE (HLA-AR) (HLA-12.4).
 GN HLA-H OR HLAH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.

RN SEQUENCE FROM N.A.
 RX MEDLINE: 82151002.
 RA MALISSEN M., MALISSEN B., JORDAN B.R.;
 RT "Exon/Intron organization and complete nucleotide sequence of an HLA
 gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:893-897(1982).
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM. COULD BE THE PRODUCT OF A PSEUDOGENE.
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J00191; AAA36218.1; ALT_INIT.
 DR PIR: A02189; HLH012.
 DR HSSP: P03989; ILSA.
 DR MIM: 142800;
 DR PROSITE: PS00290; IG_MHC; 1.
 DR PFAM: PF00047; Ig; 1.
 DR PFAM: PF00129; MHC_I; 1.
 KW MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT TRANSMEM 309 332
 FT DOMAIN 333 362
 FT CARBOHYD 110 110
 FT DISULFID 227 283
 SQ SEQUENCE 362 AA; 40850 MW; 5610F63 CRC32;

Query Match 100.0%; Score 49; DB 1; Length 362;
 Best Local Similarity 100.0%; Pred. No. 0.0037;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RENLRIALRY 10
 Db 99 RENLRIALRY 108

RESULT 15
 1A23_HUMAN
 ID 1A23_HUMAN STANDARD; PRT; 365 AA.
 AC P30447;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-23(A-9) ALPHA CHAIN
 DE PRECURSOR.
 DE HLA-A OR HLA-A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A. (A*2301).
 RX MEDLINE: 92104637.
 RA LITTLE A.-M., MADRIGAL J.A., PARHAM P.;
 RT "Molecular definition of an elusive third HLA-A9 molecule: HLA-A9.3.";
 RL Immunogenetics 35:41-45(1992).
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-

CC MICROGLOBULIN).
 CC -!- POLYMORPHISM: THE ONLY ALLELE OF A-23 KNOWN IS A*2301 WHICH IS
 CC SHOWN HERE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M64742; AAA03662.1;
 DR HSSP: P01892; IAQD.
 DR MIM: 142800;
 DR PROSITE: PS00290; IG_MHC; 1.
 DR PFAM: PF00047; Ig; 1.
 DR PFAM: PF00129; MHC_I; 1.
 KW MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 365
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT TRANSMEM 309 332
 FT DOMAIN 333 365
 FT CARBOHYD 110 110
 FT DISULFID 125 188
 FT DISULFID 227 283
 SQ SEQUENCE 365 AA; 40732 MW; BIC21094 CRC32;
 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
 A-23(A-9) ALPHA CHAIN.
 EXTRACELLULAR ALPHA-1.
 EXTRACELLULAR ALPHA-2.
 EXTRACELLULAR ALPHA-3.
 CONNECTING PEPTIDE.
 CYTOPLASMIC TAIL.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 Query Match 100.0%; Score 49; DB 1; Length 365;
 Best Local Similarity 100.0%; Pred. No. 0.0037;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RENLRIALRY 10
 Db 99 RENLRIALRY 108

Search completed: February 8, 2000, 00:59:49
 Job time: 3778 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 13:17:35 ; Search time 209.03 seconds
(without alignments)
3.317 Million cell updates/sec

Title: US-08-653-294-8

Perfect score: 49

Sequence: 1 RENLIRALRY 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

SPTREMBL_12.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 49 | 100.0 | 89 | 7 019569 | 019569 homo sapien |
| 2 | 49 | 100.0 | 90 | 7 046697 | 046697 gorilla gor |
| 3 | 49 | 100.0 | 133 | 7 019189 | 019189 homo sapien |
| 4 | 49 | 100.0 | 138 | 7 078209 | 078209 homo sapien |
| 5 | 49 | 100.0 | 172 | 7 019770 | 019770 homo sapien |
| 6 | 49 | 100.0 | 172 | 7 019774 | 019774 homo sapien |
| 7 | 49 | 100.0 | 172 | 7 019775 | 019775 homo sapien |
| 8 | 49 | 100.0 | 172 | 7 019780 | 019780 homo sapien |
| 9 | 49 | 100.0 | 172 | 7 095364 | 095364 homo sapien |
| 10 | 49 | 100.0 | 172 | 7 019771 | 019771 homo sapien |
| 11 | 49 | 100.0 | 172 | 7 019772 | 019772 homo sapien |
| 12 | 49 | 100.0 | 172 | 7 019773 | 019773 homo sapien |
| 13 | 49 | 100.0 | 175 | 7 029694 | 029694 homo sapien |
| 14 | 49 | 100.0 | 180 | 7 019607 | 019607 homo sapien |
| 15 | 49 | 100.0 | 180 | 7 019608 | 019608 homo sapien |
| 16 | 49 | 100.0 | 180 | 7 019609 | 019609 homo sapien |
| 17 | 49 | 100.0 | 180 | 7 019610 | 019610 homo sapien |
| 18 | 49 | 100.0 | 180 | 7 019611 | 019611 homo sapien |
| 19 | 49 | 100.0 | 180 | 7 019612 | 019612 homo sapien |
| 20 | 49 | 100.0 | 180 | 7 019613 | 019613 homo sapien |

| | | | | | |
|----|----|-------|-----|----------|--------------------|
| 21 | 49 | 100.0 | 181 | 7 046703 | 046703 homo sapien |
| 22 | 49 | 100.0 | 181 | 7 062917 | 062917 homo sapien |
| 23 | 49 | 100.0 | 181 | 7 062892 | 062892 homo sapien |
| 24 | 49 | 100.0 | 181 | 7 062899 | 062899 homo sapien |
| 25 | 49 | 100.0 | 181 | 7 062920 | 062920 homo sapien |
| 26 | 49 | 100.0 | 181 | 7 062922 | 062922 homo sapien |
| 27 | 49 | 100.0 | 181 | 7 062923 | 062923 homo sapien |
| 28 | 49 | 100.0 | 181 | 7 019623 | 019623 homo sapien |
| 29 | 49 | 100.0 | 181 | 7 019747 | 019747 homo sapien |
| 30 | 49 | 100.0 | 181 | 7 029667 | 029667 homo sapien |
| 31 | 49 | 100.0 | 181 | 7 030198 | 030198 homo sapien |
| 32 | 49 | 100.0 | 181 | 7 029708 | 029708 homo sapien |
| 33 | 49 | 100.0 | 181 | 7 019631 | 019631 homo sapien |
| 34 | 49 | 100.0 | 181 | 7 019769 | 019769 homo sapien |
| 35 | 49 | 100.0 | 181 | 7 029724 | 029724 homo sapien |
| 36 | 49 | 100.0 | 181 | 7 029910 | 029910 homo sapien |
| 37 | 49 | 100.0 | 181 | 7 079559 | 079559 homo sapien |
| 38 | 49 | 100.0 | 181 | 7 029679 | 029679 homo sapien |
| 39 | 49 | 100.0 | 181 | 7 019521 | 019521 homo sapien |
| 40 | 49 | 100.0 | 181 | 7 019597 | 019597 homo sapien |
| 41 | 49 | 100.0 | 181 | 7 029909 | 029909 homo sapien |
| 42 | 49 | 100.0 | 181 | 7 029701 | 029701 homo sapien |
| 43 | 49 | 100.0 | 181 | 7 029841 | 029841 homo sapien |
| 44 | 49 | 100.0 | 181 | 7 019354 | 019354 gorilla gor |
| 45 | 49 | 100.0 | 181 | 7 029765 | 029765 homo sapien |

ALIGNMENTS

RESULT 1
019569 PRELIMINARY; PRT; 89 AA.
ID 019569
AC 019569;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE MHC CLASS I ANTIGEN (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CAO K., BURDETT L., ZHANG G., FERNANDEZ-VINA M.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF017320; AAB70286.2; -.
KW MHC.
FT NON_TER 1
FT NON_TER 89
SQ SEQUENCE 89 AA; 10606 MW; 99D11089 CRC32;

Query Match 100.0%; Score 49; DB 7; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLIRALRY 10
Db 74 RENLIRALRY 83

RESULT 2
046697 PRELIMINARY; PRT; 90 AA.
ID 046697
AC 046697;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DE MHC CLASS I ANTIGEN HLA-H ORTHOLOG (FRAGMENT).
GN HLA-H.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Homnidae; Gorilla.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SHAMBIA;
 RA GRIMSLEY C., MATHER K.A., OBER C.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF022172; AAC99794.1; -
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 90 90
 SQ SEQUENCE 90 AA; 10689 MW; 5E5F2495 CRC32;

Query Match 100.0%; Score 49; DB 7; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.01;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLIRALRY 10
 |||||
 Db 75 RENLIRALRY 84

RESULT 3
 OI9189 PRELIMINARY; PRT; 133 AA.
 AC OI9189;
 DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
 DE MHC CLASS I HISTOCOMPATIBILITY ANTIGEN-B (HLA-B-27KSH) (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LEUKOCYTE;
 RA PETERSDORF E.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U18659; AAB60357.1; -
 DR MIM; 142830; -
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC I.
 FT NON_TER 1 1
 FT NON_TER 133 133
 SQ SEQUENCE 133 AA; 15491 MW; 3A3BC802 CRC32;

Query Match 100.0%; Score 49; DB 7; Length 133;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLIRALRY 10
 |||||
 Db 27 RENLIRALRY 36

RESULT 4
 O78209 PRELIMINARY; PRT; 138 AA.
 AC O78209;
 DT 01-NOV-1998 (TEMBLrel. 08, Created)
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
 DT 01-MAY-1999 (TEMBLrel. 10, Last annotation update)
 DE HUMAN LEUKOCYTE ANTIGEN PRECURSOR (FRAGMENT).
 GN HLA-A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 9800772.

RA LAFORET M., FROELICH N., PARISSIADIS A., BAUSINGER H., PFEIFFER B.,
 RA TONGIO M.M.;
 RT "an intronic mutation responsible for a low level of expression of an
 RT HLA-A*24 allele.";
 RL Tissue Antigens 50:340-346(1997).
 DR EMBL; Z72423; CAA96533.1; -
 DR PFAM; PF00129; MHC_I; 1.
 KW Signal; MHC.
 FT SIGNAL 1 24
 FT NON_TER 138 138
 SQ SEQUENCE 138 AA; 15610 MW; B8417FA0 CRC32;

Query Match 100.0%; Score 49; DB 7; Length 138;
 Best Local Similarity 100.0%; Pred. No. 0.016;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLIRALRY 10
 |||||
 Db 99 RENLIRALRY 108

RESULT 5
 OI9770 PRELIMINARY; PRT; 172 AA.
 AC OI9770;
 DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE MHC CLASS I HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGRONG E.,
 RA BEJCHANDRA S., JUJI T., TOKUNAGA K.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U90421; AAB50144.1; -
 DR EMBL; U90420; AAB50144.1; JOINED.
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 172 172
 SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 100.0%; Score 49; DB 7; Length 172;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLIRALRY 10
 |||||
 Db 68 RENLIRALRY 77

RESULT 6
 OI9774 PRELIMINARY; PRT; 172 AA.
 AC OI9774;
 DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE MHC CLASS I HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGRONG E.,
 RA BEJCHANDRA S., BLASCZYK R., GROSSE-WILDE H.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U0423; AAB50145.1; -
DR EMBL; U0422; AAB50145.1; JOINED.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.

FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 100.0%; Score 49; DB 7; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RENLIALRY 10
DB 68 RENLIALRY 77

RESULT 7
O19775 PRELIMINARY; PRT; 172 AA.
AC O19775;

DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)

DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,
RA BEUCHANDRA S.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U0425; AAB50146.1; -
DR EMBL; U0424; AAB50146.1; JOINED.

DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 100.0%; Score 49; DB 7; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RENLIALRY 10
DB 68 RENLIALRY 77

RESULT 8
O19780 PRELIMINARY; PRT; 172 AA.
AC O19780;

DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)

DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,
RA BEUCHANDRA S., JUJI T., TOKUNAGA K.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U0419; AAB50143.1; -
DR EMBL; U0418; AAB50143.1; JOINED.

DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 100.0%; Score 49; DB 7; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RENLIALRY 10
DB 68 RENLIALRY 77

RESULT 9
O95364 PRELIMINARY; PRT; 172 AA.
AC O95364;

DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)

DE MHC HLA-B*51 PROTEIN (FRAGMENT).
GN HLA-B*51FA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]
RP SEQUENCE FROM N.A.
RA BLASCZYK R.;

RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X96473; CAA65327.1; -
DR PFAM; PF00129; MHC_I; 1.
KW MHC.

FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 19942 MW; 1A73E47D CRC32;

Query Match 100.0%; Score 49; DB 7; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RENLIALRY 10
DB 65 RENLIALRY 74

RESULT 10
O19771 PRELIMINARY; PRT; 172 AA.
AC O19771;

DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)

DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D.,
RA SIRIBOONRIT U., RUNGROUNG E., BEUCHANDRA S.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U0612; AAB50151.1; -
DR EMBL; U0611; AAB50151.1; JOINED.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 20026 MW; 4D9A1043 CRC32;

Query Match 100.0%; Score 49; DB 7; Length 172;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRIALRY 10
 Db 68 RENLRIALRY 77

RESULT 11
 O19772 PRELIMINARY; PRT; 172 AA.

AC O19772; (Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE MHC CLASS I HLA-B (FRAGMENT).
 GN HLA-B.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

RP SEQUENCE FROM N.A.
 RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D.,
 RA SRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U90614; AAB50244.1; JOINED.
 DR EMBL; U90613; AAB50244.1; JOINED.
 DR PFAM; PF00129; MHC_I; 1.

KW MHC.

FT NON_TER 172 172

FT NON_TER 172 172

SQ SEQUENCE 172 AA; 20026 MW; 4D9A1043 CRC32;

Query Match 100.0%; Score 49; DB 7; Length 172;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRIALRY 10
 Db 68 RENLRIALRY 77

RESULT 12

O19773 PRELIMINARY; PRT; 172 AA.

AC O19773; (Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE MHC CLASS I HLA-B (FRAGMENT).
 GN HLA-B.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

RP SEQUENCE FROM N.A.

RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D.,
 RA SRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U90616; AAB50245.1; JOINED.
 DR EMBL; U90615; AAB50245.1; JOINED.

DR PFAM; PF00129; MHC_I; 1.

KW MHC.

FT NON_TER 172 172

FT NON_TER 172 172

SQ SEQUENCE 172 AA; 20052 MW; F6214671 CRC32;

Query Match 100.0%; Score 49; DB 7; Length 172;

Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRIALRY 10
 Db 68 RENLRIALRY 77

RESULT 13

Q29694 PRELIMINARY; PRT; 175 AA.

AC Q29694; (Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE MHC CLASS I HLA-B ANTIGEN (FRAGMENT).
 GN HLA-B.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

RP SEQUENCE FROM N.A.

RA PETERSDORF E.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U28759; AAB60367.1; JOINED.

DR HSSP; P10318; 1ROG.

DR PFAM; PF00129; MHC_I; 1.

KW MHC.

FT NON_TER 1 1

FT VARIANT 67 67 Y -> D.

FT VARIANT 73 73 I -> T.

FT NON_TER 175 175

SQ SEQUENCE 175 AA; 20332 MW; 83A0C5C3 CRC32;

Query Match 100.0%; Score 49; DB 7; Length 175;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRIALRY 10
 Db 68 RENLRIALRY 77

RESULT 14

O19607 PRELIMINARY; PRT; 180 AA.

AC O19607; (Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE MHC CLASS I HLA-A (FRAGMENT).
 GN HLA-A.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

RP SEQUENCE FROM N.A.

RA CHANDANAYINGYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
 RA RUNGROUNG E., BEJCHANDRA S.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF030920; AAB87056.1; JOINED.

DR EMBL; AF030919; AAB87056.1; JOINED.

DR HSSP; P01891; 1TMC.

DR PFAM; PF00129; MHC_I; 1.

KW MHC.

FT NON_TER 1 1

FT NON_TER 180 180

SQ SEQUENCE 180 AA; 20811 MW; CECC3537 CRC32;

Query Match 100.0%; Score 49; DB 7; Length 180;

Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RNLRLRLRY 10
| | | | |
Db 74 RNLRLRLRY 83

RESULT 15

O19608 PRELIMINARY; PRT; 180 AA.
AC O19608;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DE 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE MHC CLASS I HLA-A (FRAGMENT).
GN HLA-A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYINGTONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
RA RUNGROUNG E., BEJCHANDRA S.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF030922; AAB87057.1; "-"
DR EMBL; AF030921; AAB87057.1; JOINED.
DR HSSP; P01891; ITMC.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 180
SQ SEQUENCE 180 AA; 20811 MW; CECC3537 CRC32;

Query Match 100.0%; Score 49; DB 7; Length 180;

Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RNLRLRLRY 10
| | | | |
Db 74 RNLRLRLRY 83

Search completed: February 8, 2000, 13:17:35
Job time: 32484 sec

THIS PAGE BLANK (USPTO)
THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-8 to: GenEmbl.* out_format : pfs
 Date: Feb 8, 2000 4:37 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=framet_p2n.model -DEV=xlp
 -Q/cnrl_1/USPto.spool/US08653294/runat_04022000_160701_15779/app_query.fasta.1
 -DB=GenEmbl -QFMT=fastap -SUFFIX=ige -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -CGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEX=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -ALIGN=15 -MODE=LOCAL
 -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294
 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-8
 Query length: 10
 Database: GenEmbl.*
 Database sequences: 821193
 Database length: 1518192014
 Search time (sec): 11370.480000

score_list:

| Sequence | Strd Orig | ZScore | Escore | Len | Documentation |
|-------------------|-----------|--------|--------|-----|----------------------------------|
| gb_pr1.HSHLABHBA | + | 49.00 | 0.0103 | 250 | Y08692 H.sapiens HLA-B gene, ex |
| gb_pr1.HSHLABHBB | + | 49.00 | 0.0103 | 250 | Y08693 H.sapiens HLA-B gene, ex |
| gb_pr1.HSHLABHBD | + | 49.00 | 0.0103 | 250 | Y08694 H.sapiens HLA-B gene, ex |
| gb_pr2.HSHLABB1 | + | 49.00 | 0.0103 | 250 | U090420 Human cell line THAI DCH |
| gb_pr2.HSHLABD1 | + | 49.00 | 0.0103 | 250 | U090611 Human cell line THAI DCH |
| gb_pr2.HSHLABF1 | + | 49.00 | 0.0103 | 250 | U090613 Human cell line THAI DCH |
| gb_pr2.HSHLABG1 | + | 49.00 | 0.0103 | 250 | U090615 Human cell line THAI DCH |
| gb_pr2.HSHLABI1 | + | 49.00 | 0.0103 | 250 | U090422 Human cell line THAI DCH |
| gb_pr2.HSHLABJ1 | + | 49.00 | 0.0103 | 250 | U090424 Human cell line THAI DCH |
| gb_pr2.HSHLABK1 | + | 49.00 | 0.0103 | 250 | U090418 Human cell line THAI DCH |
| gb_pr4.AFO22159 | + | 49.00 | 0.0105 | 255 | AF022159 Homo sapiens isolate C |
| gb_pr4.AFO22160 | + | 49.00 | 0.0107 | 259 | AF022160 Homo sapiens isolate C |
| gb_pat.I14590 | + | 49.00 | 0.0112 | 270 | I14590 Sequence 67 from patent |
| gb_pat.I14591 | + | 49.00 | 0.0112 | 270 | I14591 Sequence 68 from patent |
| gb_pat.I14592 | + | 49.00 | 0.0112 | 270 | I14592 Sequence 69 from patent |
| gb_pat.I14593 | + | 49.00 | 0.0112 | 270 | I14593 Sequence 70 from patent |
| gb_pr1.HS51EX2 | + | 49.00 | 0.0112 | 270 | Y11228 H.sapiens HLA-B*51 gene, |
| gb_pr1.HSHLAL1 | + | 49.00 | 0.0112 | 270 | X82161 H.sapiens HLA-A gene, al |
| gb_pr1.HUMHB512 | + | 49.00 | 0.0112 | 270 | M2787 Human MHC class I HLA-B* |
| gb_pr1.HUMHBW2 | + | 49.00 | 0.0112 | 270 | M2794 Human MHC class I HLA-B* |
| gb_pr2.HMCHL1 | + | 49.00 | 0.0112 | 270 | U63395 Human MHC class I antige |
| gb_pr2.HSA011699 | + | 49.00 | 0.0112 | 270 | AJ011699 Homo sapiens HLA-A gen |
| gb_pr2.HSA133780 | + | 49.00 | 0.0112 | 270 | AJ133780 Homo sapiens HLA-B gen |
| gb_pr2.HSA238971 | + | 49.00 | 0.0112 | 270 | AJ238971 Homo sapiens HLA-B gen |
| gb_pr2.HSA239035 | + | 49.00 | 0.0112 | 270 | AJ239035 Homo sapiens HLA-A gen |
| gb_pr2.HSHL24J01 | + | 49.00 | 0.0112 | 270 | U37110 Human HLA-A*24 gene, alle |
| gb_pr2.HSHL24SA01 | + | 49.00 | 0.0112 | 270 | U37114 Human HLA-A*24 gene, alle |
| gb_pr2.HSHL24Y01 | + | 49.00 | 0.0112 | 270 | U37112 Human HLA-A*24 gene, alle |
| gb_pr2.HSHL24S1 | + | 49.00 | 0.0112 | 270 | U18987 Human MHC class I antige |
| gb_pr2.HSHLABEXN1 | + | 49.00 | 0.0112 | 270 | U76400 Human HLA-B gene, allele |
| gb_pr2.HUMHHLAFA | + | 49.00 | 0.0112 | 270 | L43530 Homo sapiens (clone K92) |
| gb_pr2.L43528 | + | 49.00 | 0.0112 | 270 | L43528 Homo sapiens (clone K62) |
| gb_pr2.X83402 | + | 49.00 | 0.0112 | 270 | X83402 H.sapiens HLA-B gene, ex |
| gb_pr3.HS1109HLA1 | + | 49.00 | 0.0112 | 270 | AF030923 Homo sapiens MHC class |
| gb_pr3.HS1109HLA1 | + | 49.00 | 0.0112 | 270 | AF030931 Homo sapiens MHC class |
| gb_pr3.HS2402HLA1 | + | 49.00 | 0.0112 | 270 | AF067436 Homo sapiens MHC class |
| gb_pr3.HS507HLA1 | + | 49.00 | 0.0112 | 270 | AF030921 Homo sapiens MHC class |
| gb_pr3.HS5108HLA1 | + | 49.00 | 0.0112 | 270 | U52815 Homo sapiens class I and |
| gb_pr3.HS524HLA1 | + | 49.00 | 0.0112 | 270 | AF030919 Homo sapiens MHC class |
| gb_pr3.HS5342HLA1 | + | 49.00 | 0.0112 | 270 | AF030925 Homo sapiens MHC class |
| gb_pr3.HS538HLA1 | + | 49.00 | 0.0112 | 270 | AF030911 Homo sapiens MHC class |
| gb_pr3.HS5802HLA1 | + | 49.00 | 0.0112 | 270 | U52813 Human MHC class I antige |

gb_pr3.HS611HLA1 + 49.00 192.10 0.0112 270 AF030927 Homo sapiens MHC cl
 gb_pr3.HS639HLA1 + 49.00 192.10 0.0112 270 AF030929 Homo sapiens MHC cl
 gb_pr4.HSKM315S1 + 49.00 192.10 0.0112 270 U59699 Human MHC class I (HL

seq_name: gb_pr1.HSHLABHBA

seq_documentation_block:
 LOCUS HSHLABHBA 250 bp DNA PRI 10-OCT-1996
 DEFINITION H.sapiens HLA-B gene, exon 2, HB(a) allele.
 ACCESSION Y08692
 VERSION Y08692.1 GI:1619287
 KEYWORDS HLA-B gene; human leukocyte antigen.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS Rojas-Munoz,A., Mendez,I. and Yunis,I.
 TITLE Molecular evolution of HLA-B locus in a small population amerindian
 community :The Nukak-Maku

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 250)
 AUTHORS Rojas-Munoz,A.
 TITLE Direct Submission
 JOURNAL Submitted (07-OCT-1996) A. Rojas-Munoz, National Institute Of
 Health, Immunogenetics, Avenida El Dorado, Carrera 50, Santafe De
 Bogota / Zona 6, COLOMBIA

FEATURES
 Location/Qualifiers
 1..250

/organism="Homo sapiens"
 /isolate="Nancy-28"
 /isolate="from amerindian community Nukak-Maku"
 /db_xref="taxon:9606"
 /chromosome="6"
 /dev_stage="adult"
 /tissue_type="blood"
 /cell_type="white"
 /lab_host="E.coli TGI"
 /clone="CHB1(a)"
 /clone="CHB2(A)"
 /clone="CHB4(a)"
 14..250
 /gene="HLA-B"
 /gene="HLA-B"
 /note="allel HB(a)"
 /number=2
 BASE COUNT 54 a 78 c 85 g 33 t
 ORIGIN

alignment_scores:
 Quality: 49.00 Length: 10
 Ratio: 4.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-8 x HSHLABHBA ..
 Align seg 1/1 to: HSHLABHBA from: 1 to: 250
 1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
 |||||
 209 CGAGAACCTCGCGATCGCGCTCCGCTAC 238

seq_name: gb_pr1.HSHLABHBB

seq_documentation_block:
 LOCUS HSHLABHBB 250 bp DNA PRI 10-OCT-1996
 DEFINITION H.sapiens HLA-B gene, exon 2, HB(b) allele.
 ACCESSION Y08693
 VERSION Y08693.1 GI:1619288
 KEYWORDS HLA-B gene; human leukocyte antigen.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS Rojas-Munoz, A., Mendez, I. and Yunis, I.
 TITLE Molecular evolution of HLA-B locus in a small population amerindian community : The Nukak-Maku
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 250)
 AUTHORS Rojas-Munoz, A.
 TITLE Direct Submission
 JOURNAL Submitted (07-OCT-1996) A. Rojas-Munoz, National Institute Of Health, Immunogenetics, Avenida El Dorado, Carrera 50, Santafe De Bogota / Zona 6, COLOMBIA
 FEATURES source
 1. .250
 /organism="Homo sapiens"
 /isolate="isabel-26"
 /isolate="from amerindian community Nukak-Maku"
 /db_xref="taxon:9606"
 /chromosome="6"
 /dev_stage="adult"
 /tissue_type="blood"
 /cell_type="white"
 /lab_host="E.coli TGI"
 /clone="CHC2(b)"
 /clone="CHC3(b)"
 /clone="CHC4(b)"
 14. .250
 /gene="HLA-B"
 <14. .>250
 /gene="HLA-B"
 /note="allele HB(b)"
 /number=2
 BASE COUNT 51 a 78 c 87 g 34 t
 ORIGIN
 alignment_scores
 Quality: 49.00 Length: 10
 Ratio: 4.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block
 US-08-653-294-8 x HSLABHBB ..
 Align seg 1/1 to: HSLABHBB from: 1 to: 250
 1 ArgGlusnLeuArglleAlaLeuArgTyr 10
 |||||
 209 CGAGAGAACCTGGGATCGGCTCGCTAC 238
 seq_name: gb_pr1:HSLABHBD
 seq_documentation_block
 LOCUS HSLABHBD 250 bp DNA PRI 10-OCT-1996
 DEFINITION H.sapiens HLA-B gene, exon 2, HB(d) allele.
 ACCESSION Y08694
 VERSION Y08694.1 GI:1619289
 KEYWORDS HLA-B gene; human leukocyte antigen.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS Rojas-Munoz, A., Mendez, I. and Yunis, I.
 TITLE Molecular evolution of HLA-B locus in a small population amerindian community : The Nukak-Maku
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 250)
 AUTHORS Rojas-Munoz, A.
 TITLE Direct Submission

JOURNAL Submitted (07-OCT-1996) A. Rojas-Munoz, National Institute Of Health, Immunogenetics, Avenida El Dorado, Carrera 50, Santafe De Bogota / Zona 6, COLOMBIA
 FEATURES source
 1. .250
 /organism="Homo sapiens"
 /isolate="Norman-51"
 /isolate="from amerindian community Nukak-Maku"
 /db_xref="taxon:9606"
 /chromosome="6"
 /dev_stage="adult"
 /tissue_type="blood"
 /cell_type="white"
 /lab_host="E.coli TGI"
 /clone="CHC1(d)"
 14. .250
 /gene="HLA-B"
 <14. .>250
 /gene="HLA-B"
 /note="allele HB(d)"
 /number=2
 BASE COUNT 58 a 78 c 79 g 35 t
 ORIGIN
 alignment_scores
 Quality: 49.00 Length: 10
 Ratio: 4.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block
 US-08-653-294-8 x HSLABHBD ..
 Align seg 1/1 to: HSLABHBD from: 1 to: 250
 1 ArgGlusnLeuArglleAlaLeuArgTyr 10
 |||||
 209 CGAGAGAACCTGGGATCGGCTCGCTAC 238
 seq_name: gb_pr2:HSLABBI
 seq_documentation_block
 LOCUS HSLABBI 250 bp DNA PRI 22-MAR-1997
 DEFINITION Human cell line THAI DCHO10 MHC class I HLA-B gene (allele HLA-B*1513), exon 2.
 ACCESSION U90420
 VERSION U90420.1 GI:1905830
 KEYWORDS
 SEGMENT 1 of 2
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS Chandanayingyong, D., Sirikong, M., Longta, K., Srinak, D., Rungroung, E., Bejchandra, S., Juji, T. and Tokunaga, K.
 TITLE B15 alleles (B*1513)
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 250)
 AUTHORS Chandanayingyong, D., Sirikong, M., Longta, K., Srinak, D., Rungroung, E., Bejchandra, S., Juji, T. and Tokunaga, K.
 TITLE Direct Submission
 JOURNAL Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine, Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700, Thailand
 FEATURES source
 1. .250
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="VI"
 /map="6p21"
 /cell_type="lymphoblastoid"
 /cell_line="THAI DCHO10"

```
exon
1..250
/gene="HLA-B"
/number=2
/product="MHC class I HLA-B"
55 a 83 c 80 g 32 t

alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-08-653-294-8 x HSHLABB1
..
Align seg 1/1 to: HSHLABB1 from: 1 to: 250
1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||
202 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 231

seq_name: gb_pr2:HSHLABD1

seq_documentation_block: 250 bp DNA PRI 22-MAR-1997
LOCUS HSHLABD1
DEFINITION Human cell line THAI DCH010 MHC class I HLA-B gene (allele
HLA-B*51V), exon 2.
ACCESSION U90611
VERSION U90611.1 GI:1905865
KEYWORDS
SEGMENT
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 250)
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
B*51V alleles
Unpublished
JOURNAL
2 (bases 1 to 250)
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
Direct Submission
TITLE
JOURNAL
Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES
Source
1..250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/haplotype="(b)HLA-A2,B77,Cw8N,DR12(DR52),DQ7/(c)A11.1,
B51V,Cw14,DR12(DR52),DQ1"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH010"
1..250
/gene="HLA-B"
/notes="Allele: HLA-B*51V; similar to exon 2 of B*5105"
/number=2
/product="MHC class I HLA-B"
56 a 82 c 80 g 32 t

BASE COUNT
ORIGIN

alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-08-653-294-8 x HSHLABD1
..
Align seg 1/1 to: HSHLABD1 from: 1 to: 250
1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||
202 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 231

seq_name: gb_pr2:HSHLABD1

seq_documentation_block: 250 bp DNA PRI 22-MAR-1997
LOCUS HSHLABD1
DEFINITION Human cell line THAI DCH010 MHC class I HLA-B gene (allele
HLA-B*51V), exon 2.
ACCESSION U90611
VERSION U90611.1 GI:1905865
KEYWORDS
SEGMENT
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 250)
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
B*51V alleles
Unpublished
JOURNAL
2 (bases 1 to 250)
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
Direct Submission
TITLE
JOURNAL
Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES
Source
1..250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/haplotype="(b)HLA-A2,B77,Cw8N,DR12(DR52),DQ7/(c)A11.1,
B51V,Cw14,DR12(DR52),DQ1"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH010"
1..250
/gene="HLA-B"
/notes="Allele: HLA-B*51V; similar to exon 2 of B*5105"
/number=2
/product="MHC class I HLA-B"
56 a 82 c 80 g 32 t

BASE COUNT
ORIGIN

alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-08-653-294-8 x HSHLABD1
..
Align seg 1/1 to: HSHLABD1 from: 1 to: 250
1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||
202 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 231

seq_name: gb_pr2:HSHLABD1

seq_documentation_block: 250 bp DNA PRI 25-MAR-1997
LOCUS HSHLABD1
DEFINITION Human cell line THAI DCH028 MHC class I HLA-B gene (allele
HLA-B*51V), exon 2.
ACCESSION U90613
VERSION U90613.1 GI:1906033
KEYWORDS
SEGMENT
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 250)
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
B*51V alleles
Unpublished
JOURNAL
2 (bases 1 to 250)
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
Direct Submission
TITLE
JOURNAL
Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES
Source
1..250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/haplotype="(a)HLA-A1,B51V,Cw14,DR6(DR52),DQ1/(b)A2,B77,
Cw8N,DR4(DR53),DQ4"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH028"
1..250
/gene="HLA-B"
/notes="Allele: HLA-B*51V; similar to exon 2 of B*5105"
/number=2
/product="MHC class I HLA-B"
56 a 82 c 80 g 32 t

BASE COUNT
ORIGIN

alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-08-653-294-8 x HSHLABF1
..
Align seg 1/1 to: HSHLABF1 from: 1 to: 250
1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||
202 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 231

seq_name: gb_pr2:HSHLABG1

seq_documentation_block: 250 bp DNA PRI 25-MAR-1997
LOCUS HSHLABG1
DEFINITION Human cell line THAI DCH011 MHC class I HLA-B gene (allele
HLA-B*51V), exon 2.
ACCESSION U90615
```

```
US-08-653-294-8 x HSHLABD1
..
Align seg 1/1 to: HSHLABD1 from: 1 to: 250
1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||
202 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 231

seq_name: gb_pr2:HSHLABF1

seq_documentation_block: 250 bp DNA PRI 25-MAR-1997
LOCUS HSHLABF1
DEFINITION Human cell line THAI DCH028 MHC class I HLA-B gene (allele
HLA-B*51V), exon 2.
ACCESSION U90613
VERSION U90613.1 GI:1906033
KEYWORDS
SEGMENT
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 250)
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
B*51V alleles
Unpublished
JOURNAL
2 (bases 1 to 250)
Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
Siriboonrit,U., Rungroung,E. and Bejchandra,S.
Direct Submission
TITLE
JOURNAL
Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,
Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
Thailand
FEATURES
Source
1..250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/haplotype="(a)HLA-A1,B51V,Cw14,DR6(DR52),DQ1/(b)A2,B77,
Cw8N,DR4(DR53),DQ4"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH028"
1..250
/gene="HLA-B"
/notes="Allele: HLA-B*51V; similar to exon 2 of B*5105"
/number=2
/product="MHC class I HLA-B"
56 a 82 c 80 g 32 t

BASE COUNT
ORIGIN

alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-08-653-294-8 x HSHLABF1
..
Align seg 1/1 to: HSHLABF1 from: 1 to: 250
1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||
202 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 231

seq_name: gb_pr2:HSHLABF1

seq_documentation_block: 250 bp DNA PRI 25-MAR-1997
LOCUS HSHLABG1
DEFINITION Human cell line THAI DCH011 MHC class I HLA-B gene (allele
HLA-B*51V), exon 2.
ACCESSION U90615
```

VERSION U90615.1 GI:1906037
 KEYWORDS 1 of 2
 SEGMENT human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS Chandanayingyong, D., Sirikong, M., Longta, K., Srinak, D.,
 Siriboonrit, U., Rungroung, E. and Bejchandra, S.
 TITLE B*51V alleles
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 250)
 AUTHORS Chandanayingyong, D., Sirikong, M., Longta, K., Srinak, D.,
 Siriboonrit, U., Rungroung, E. and Bejchandra, S.
 TITLE Direct Submission
 JOURNAL Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,
 Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
 Thailand
 FEATURES Location/Qualifiers
 source 1..250
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /map="6p21"
 /haplotype="(a)HLA-A11.1,B51V,Cw14,DR12(DR52).DQ7/(b)A24,
 B77,Cw8N,DR12(DR52)DQ7"
 /cell_type="lymphoblastoid"
 /cell_line="THAI DCH011"
 1..250
 /gene="HLA-B"
 /note="Allele: HLA-B*51V; similar to exon 2 of B*5105"
 /number=2
 exon 56 a 82 g 32 t
 BASE COUNT 56 a 82 g 32 t
 ORIGIN
 alignment_scores:
 Quality: 49.00 Length: 10
 Ratio: 4.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-08-653-294-8 x HSHLABG1 ..
 Align seg 1/1 to: HSHLABG1 from: 1 to: 250
 1 ArgGluAsnLeuArgileAlaLeuArgTyr 10
 |||||
 202 CGAGAGACCTCGGATCGCGTCCGCTAC 231
 seq_name: gb_pr2: HSHLABG1
 seq_documentation_block:
 LOCUS HSHLABG1 250 bp DNA PRI 22-MAR-1997
 DEFINITION Human cell line THAI DCH011 MHC class I HLA-B gene (allele
 HLA-B*1513), exon 2.
 ACCESSION U90422
 VERSION U90422.1 GI:1905834
 KEYWORDS 1 of 2
 SEGMENT human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS Chandanayingyong, D., Sirikong, M., Longta, K., Srinak, D.,
 Rungroung, E., Bejchandra, S., Blasczyk, R. and Grosse-Wilde, H.
 TITLE B15 alleles (B*1513)
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 250)
 AUTHORS Chandanayingyong, D., Sirikong, M., Longta, K., Srinak, D.,
 Rungroung, E., Bejchandra, S., Blasczyk, R. and Grosse-Wilde, H.

TITLE Direct Submission
 JOURNAL Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine,
 Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
 Thailand
 FEATURES Location/Qualifiers
 source 1..250
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="VI"
 /map="6p21"
 /cell_type="lymphoblastoid"
 /cell_line="THAI DCH011"
 1..250
 /gene="HLA-B"
 /note="Allele: HLA-B*1513"
 /number=2
 exon 55 a 83 c 32 t
 BASE COUNT 55 a 83 c 32 t
 ORIGIN
 alignment_scores:
 Quality: 49.00 Length: 10
 Ratio: 4.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-08-653-294-8 x HSHLABI1 ..
 Align seg 1/1 to: HSHLABI1 from: 1 to: 250
 1 ArgGluAsnLeuArgileAlaLeuArgTyr 10
 |||||
 202 CGAGAGACCTCGGATCGCGTCCGCTAC 231
 seq_name: gb_pr2: HSHLABJ1
 seq_documentation_block:
 LOCUS HSHLABJ1 250 bp DNA PRI 22-MAR-1997
 DEFINITION Human cell line THAI DCH028 MHC class I HLA-B gene (allele
 HLA-B*1513), exon 2.
 ACCESSION U90424
 VERSION U90424.1 GI:1905838
 KEYWORDS 1 of 2
 SEGMENT human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS Chandanayingyong, D., Sirikong, M., Longta, K., Srinak, D.,
 Rungroung, E. and Bejchandra, S.
 TITLE B15 alleles (B*1513)
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 250)
 AUTHORS Chandanayingyong, D., Sirikong, M., Longta, K., Srinak, D.,
 Rungroung, E. and Bejchandra, S.
 TITLE Direct Submission
 JOURNAL Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine,
 Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
 Thailand
 FEATURES Location/Qualifiers
 source 1..250
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="VI"
 /map="6p21"
 /cell_type="lymphoblastoid"
 /cell_line="THAI DCH028"
 1..250
 /gene="HLA-B"
 /note="Allele: HLA-B*1513"
 /number=2
 exon

```

BASE COUNT      55 a      83 c      80 g      32 t
ORIGIN
/product="MHC class I HLA-B"

alignment_scores:
  Quality:      49.00      Length:      10
  Ratio:        4.900      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-08-653-294-8 x HSHLABJ1      ..
Align seg 1/1 to: HSHLABJ1 from: 1 to: 250

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||
202 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 231

seq_name: gb_pr2:HSHLABT1

seq_documentation_block:
LOCUS      HSHLABT1      250 bp      DNA      PRI      22-MAR-1997
DEFINITION      Human cell line THAI DCH009 MHC class I HLA-B gene (allele
ACCESSION      U90418
VERSION      U90418.1 GI:1905826
KEYWORDS
SEGMENT
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
              Rungrong,E., Bejchandra,S., Juji,T. and Tokunaga,K.
TITLE      B15 alleles (B*1513)
JOURNAL      Unpublished
REFERENCE
AUTHORS      Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
              Rungrong,E., Bejchandra,S., Juji,T. and Tokunaga,K.
TITLE      Direct Submission
JOURNAL      Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine,
              Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
              Thailand
FEATURES
source
1..250
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="VI"
/map="6p21"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH009"
1..250
/gene="HLA-B"
/number=2
/product="MHC class I HLA-B"
BASE COUNT      55 a      83 c      80 g      32 t
ORIGIN
/quality="MHC class I HLA-B"

alignment_scores:
  Quality:      49.00      Length:      10
  Ratio:        4.900      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-08-653-294-8 x HSHLABT1      ..
Align seg 1/1 to: HSHLABT1 from: 1 to: 250

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||
202 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 231

```

```

|||||
202 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 231

seq_name: gb_pr4:AF022159

seq_documentation_block:
LOCUS      AF022159      255 bp      DNA      PRI      05-JAN-1999
DEFINITION      Homo sapiens isolate 026 MHC class I antigen HLA-H (HLA-H)
              pseudogene, partial sequence.
ACCESSION      AF022159
VERSION      AF022159.1 GI:2655062
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Grimsley,C., Mather,K.A. and Ober,C.
TITLE      HLA-H: a pseudogene with increased variation due to balancing
              selection at neighboring loci
JOURNAL      Mol. Biol. Evol. 15 (12), 1581-1588 (1998)
MEDLINE      99083426
REFERENCE
AUTHORS      Grimsley,C., Mather,K.A. and Ober,C.
TITLE      Direct Submission
JOURNAL      Submitted (03-SEP-1997) Fred Hutchinson Cancer Research Center,
              1100 Fairview Ave. N., M374, Seattle, WA 98109, USA
FEATURES
source
1..255
Location/Qualifiers
/organism="Homo sapiens"
/isolate="026"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
/note="African-American individual"
1..255
/gene="HLA-H"
/pseudo
/number=2
<1..>255
/gene="HLA-H"
/note="MHC class I antigen HLA-H"
/codon_start=1
/pseudo
<1..>255
/gene="HLA-H"
/pseudo
BASE COUNT      50 a      81 c      86 g      38 t
ORIGIN

alignment_scores:
  Quality:      49.00      Length:      10
  Ratio:        4.900      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-08-653-294-8 x AF022159      ..
Align seg 1/1 to: AF022159 from: 1 to: 255

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||
223 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 252

seq_name: gb_pr4:AF022160

seq_documentation_block:
LOCUS      AF022160      259 bp      DNA      PRI      05-JAN-1999
DEFINITION      Homo sapiens isolate 034 MHC class I antigen HLA-H (HLA-H)
              pseudogene, partial sequence.
ACCESSION      AF022160
VERSION      AF022160.1 GI:2655063

```

KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 259)
 AUTHORS Grimsley, C., Mather, K.A. and Ober, C.
 TITLE HLA-H: a pseudogene with increased variation due to balancing selection at neighboring loci
 JOURNAL Mol. Biol. Evol. 15 (12), 1581-1588 (1998)
 MEDLINE 99083426
 REFERENCE 2 (bases 1 to 259)
 AUTHORS Grimsley, C., Mather, K.A. and Ober, C.
 TITLE Direct Submission
 JOURNAL Submitted (03-SEP-1997) Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N., M374, Seattle, WA 98109, USA
 FEATURES
 source
 1..259
 /organism="Homo sapiens"
 /isolate="034"
 /db_xref="taxon:9606"
 /chromosome="6"
 /map="6p21.3"
 /note="African-American individual"
 1..259
 /gene="HLA-H"
 /pseudo
 /number=2
 <1..>259
 /gene="HLA-H"
 /note="MHC class I antigen HLA-H"
 /codon_start=1
 /pseudo
 <1..>259
 /gene="HLA-H"
 /pseudo
 BASE COUNT 52 a 83 c 85 g 37 t 2 others
 ORIGIN
 alignment_scores:
 Quality: 49.00 Length: 10
 Ratio: 4.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-08-653-294-8 x AF022160 ..
 Align seg 1/1 to: AF022160 from: 1 to: 259
 1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
 |||||
 223 CGAGAGACCTCGGATCGCGTCCGCTAC 252
 seq_name: gb_pat:114590
 seq_documentation_block:
 LOCUS 114590 270 bp DNA PAT 26-SEP-1995
 DEFINITION Sequence 67 from patent US 5451512.
 ACCESSION 114590
 VERSION 114590.1 GI:997073
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 270)
 AUTHORS Apple, R.J., Bugawan, T.L. and Erlich, H.A.
 TITLE Methods and reagents for HLA class I A locus DNA typing
 JOURNAL Patent: US 5451512-A 67 19-SEP-1995;
 FEATURES
 source
 1..270
 /organism="unknown"
 BASE COUNT 55 a 84 c 95 g 36 t
 ORIGIN
 alignment_scores:
 Quality: 49.00 Length: 10
 Ratio: 4.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-08-653-294-8 x I14591 ..
 Align seg 1/1 to: I14591 from: 1 to: 270
 1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
 |||||
 222 CGAGAGACCTCGGATCGCGTCCGCTAC 251
 seq_name: gb_pat:114592
 seq_documentation_block:
 LOCUS 114592 270 bp DNA PAT 26-SEP-1995
 DEFINITION Sequence 69 from patent US 5451512.
 ACCESSION 114592
 VERSION 114592.1 GI:997075
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 270)
 AUTHORS Apple, R.J., Bugawan, T.L. and Erlich, H.A.
 TITLE Methods and reagents for HLA class I A locus DNA typing
 JOURNAL Patent: US 5451512-A 69 19-SEP-1995;
 FEATURES
 source
 1..270
 /organism="unknown"
 BASE COUNT 55 a 84 c 95 g 36 t
 ORIGIN

ORIGIN
 alignment_scores:
 Quality: 49.00 Length: 10
 Ratio: 4.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-08-653-294-8 x I14590 ..
 Align seg 1/1 to: I14590 from: 1 to: 270
 1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
 |||||
 222 CGAGAGACCTCGGATCGCGTCCGCTAC 251
 seq_name: gb_pat:114591
 seq_documentation_block:
 LOCUS 114591 270 bp DNA PAT 26-SEP-1995
 DEFINITION Sequence 68 from patent US 5451512.
 ACCESSION 114591
 VERSION 114591.1 GI:997074
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 270)
 AUTHORS Apple, R.J., Bugawan, T.L. and Erlich, H.A.
 TITLE Methods and reagents for HLA class I A locus DNA typing
 JOURNAL Patent: US 5451512-A 68 19-SEP-1995;
 FEATURES
 source
 1..270
 /organism="unknown"
 BASE COUNT 55 a 84 c 95 g 36 t
 ORIGIN
 alignment_scores:
 Quality: 49.00 Length: 10
 Ratio: 4.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-08-653-294-8 x I14591 ..
 Align seg 1/1 to: I14591 from: 1 to: 270
 1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
 |||||
 222 CGAGAGACCTCGGATCGCGTCCGCTAC 251
 seq_name: gb_pat:114592
 seq_documentation_block:
 LOCUS 114592 270 bp DNA PAT 26-SEP-1995
 DEFINITION Sequence 69 from patent US 5451512.
 ACCESSION 114592
 VERSION 114592.1 GI:997075
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 270)
 AUTHORS Apple, R.J., Bugawan, T.L. and Erlich, H.A.
 TITLE Methods and reagents for HLA class I A locus DNA typing
 JOURNAL Patent: US 5451512-A 69 19-SEP-1995;
 FEATURES
 source
 1..270
 /organism="unknown"
 BASE COUNT 55 a 84 c 95 g 36 t
 ORIGIN

```
alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-08-653-294-8 x IL4592  ..
  Align seg 1/1 to: IL4592 from: 1 to: 270
      1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
        |||||||
      222 CGAGAGAACCTGGCGATCGCGCTCGGTAC 251
```

THIS PAGE BLANK (USPTO)


```

Command line parameters:
-Model=frame+pnz.model -DEV=xlp
-0/cgnl_1/USPTO_spool/US08653294/runat_04022000.160701.15807/appf
-DB=N_Genesec_36 -QFMT=fastap -SUFFIX=ring -GAPOF=12.000
-GAEXT=4.000 -MINMATCH=0.100 -LOOCPCL=0.000 -LOOPEXT=0.000
-GAPOF=4.500 -QGAPEXT=0.050 -XGAPOF=10.000 -XGAEXT=0.500
-GAPOF=6.000 -FGAPEXT=7.000 -YGAPOF=10.000 -YGAEXT=0.500
-DELOB=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=20 -THR_SCORE=pct
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT
-THREADS=1

```

```

Search information block:
Query: US-08-653-294-8
Query length: 10
Database: N_Geneseq_36.*
Database sequences: 311585
Database length: 125096042
Search time (sec): 590.520000

```

| Sequence | Strnd | Orig | zScore | EScore | Len | Documentation |
|------------------------|-------|-------|--------|---------|--------|------------------------------------|
| N_Genesed_36_Q029167 | + | 49.00 | 185.16 | 0.0103 | 270 | HLA-B*52 exon 2 alpha-1 domain |
| N_Genesed_36_Q001834 | + | 49.00 | 172.82 | 0.0515 | 1086 | Sequence encoding HLA-B*51 antigen |
| N_Genesed_36_Q001822 | + | 49.00 | 172.52 | 0.0515 | 1086 | Sequence encoding HLA-B*52 antigen |
| N_Genesed_36_Q005893 | + | 49.00 | 172.59 | 0.0517 | 1089 | HLA-B*51 gene for production of |
| N_Genesed_36_Q005701 | + | 49.00 | 172.59 | 0.0517 | 1089 | HLA-B*52 gene for production of |
| N_Genesed_36_Q012114 | + | 49.00 | 172.59 | 0.0517 | 1089 | HLA-B*53 exon. HLA-B*53 gene, |
| N_Genesed_36_K33945 | + | 35.00 | 111.98 | 122.96 | 2056 | Human HCMV Inducible gene, seq |
| N_Genesed_36_K70935 | + | 34.00 | 114.32 | 91.04 | 1026 | Sequence encoding the human hla |
| N_Genesed_36_K702225 | + | 34.00 | 102.35 | 422.82 | 3874 | Sequence of genomic DNA encoding |
| N_Genesed_36_K761639 | + | 34.00 | 97.61 | 776.25 | 6553 | HLA B*7 consensus sequence, de |
| N_Genesed_36_K73804 | + | 34.00 | 82.49 | 5.4e+03 | 35100 | ! KSHV LTR DNA (nucleotides 70 |
| N_Genesed_36_K715941 | + | 34.00 | 70.18 | 2.6e+04 | 137507 | ! KSHV long unique coding reg |
| N_Genesed_36_K718900 | + | 33.00 | 118.38 | 54.06 | 423 | Type II topoisomerase database |
| N_Genesed_36_K517732 | + | 33.00 | 110.83 | 142.42 | 978 | DNA encoding a human secreted p |
| N_Genesed_36_K800555 | + | 33.00 | 110.53 | 147.99 | 1011 | Beta-glucuronidase-contrig inser |
| N_Genesed_36_Q005916 | + | 33.00 | 110.53 | 147.99 | 1011 | Promoter sequence and N-termi |
| N_Genesed_36_K709025 | + | 33.00 | 100.53 | 533.48 | 3066 | ! Homo sapiens 20q13 amplicon |
| N_Genesed_36_Q043182 | + | 33.00 | 99.47 | 611.63 | 3451 | MTAce2 expression construct. |
| N_Genesed_36_K024977 | + | 33.00 | 95.88 | 969.32 | 5140 | DNA encoding soluble mannose |
| N_Genesed_36_Q043183 | + | 33.00 | 94.48 | 1.2e+03 | 6000 | MTAceA1 cassette for aceA and |
| N_Genesed_36_K703430 | + | 33.00 | 92.82 | 1.6e+03 | 7742 | ! Escherichia coli gus operon g |
| N_Genesed_36_K74414 | + | 33.00 | 86.92 | 3.1e+03 | 14051 | ! Staphylococcus aureus contig |
| N_Genesed_36_K20248_06 | - | 33.00 | 68.27 | 3.3e+04 | 110000 | ! Continuation (7 of 10) of |
| N_Genesed_36_K20248_06 | - | 33.00 | 68.16 | 3.3e+04 | 111309 | ! Borrelia burgdorferi polyonu |
| N_Genesed_36_K069946 | + | 32.00 | 132.22 | 9.17 | 59 | U7.6 L3' PCR primer for U7.6 v |
| N_Genesed_36_KV52995 | + | 32.00 | 130.68 | 11.18 | 70 | Oligonucleotide used in the cou |
| N_Genesed_36_K17619 | + | 32.00 | 129.47 | 13.04 | 80 | ! C111 scFV VL PCR primer 6. Singl |
| N_Genesed_36_K17619 | + | 32.00 | 129.14 | 13.61 | 83 | ! Vb3'AL2 PCR primer for U7.6 v |
| N_Genesed_36_K092441 | + | 32.00 | 112.36 | 117.00 | 534 | ! Sequence of carcinoembryonic |
| N_Genesed_36_K76507 | + | 32.00 | 104.51 | 320.22 | 1276 | Heat-resistant barley beta-amyl |
| N_Genesed_36_K19263 | + | 32.00 | 103.42 | 368.25 | 1440 | ! Staphylococcus aureus ribonuc |
| N_Genesed_36_K19262 | + | 32.00 | 101.16 | 492.24 | 1851 | ! Staphylococcus aureus ribonuc |
| N_Genesed_36_K071567 | + | 32.00 | 100.33 | 547.97 | 2031 | ! Carcinoembryonic antigen DNA |
| N_Genesed_36_K736495 | - | 32.00 | 100.20 | 556.72 | 2059 | ! Immunogenic carcinoembryonic |
| N_Genesed_36_K082807 | + | 32.00 | 100.04 | 568.61 | 2097 | ! Carcinoembryonic antigen cDNA |
| N_Genesed_36_K765072 | + | 32.00 | 99.85 | 582.73 | 2142 | Heat-resistant barley beta-amyl |
| N_Genesed_36_K733023 | + | 32.00 | 99.52 | 607.33 | 2220 | ! Carcinoembryonic antigen gene |
| N_Genesed_36_K067869 | - | 32.00 | 99.01 | 648.30 | 2349 | ! H6/CEA expression cassette fr |
| N_Genesed_36_K067868 | - | 32.00 | 98.69 | 675.49 | 2434 | ! H6/CEA expression cassette fr |
| N_Genesed_36_K81584 | - | 32.00 | 97.31 | 807.01 | 2839 | ! LV7 cDNA encoding carcinoemb |
| N_Genesed_36_N92780 | - | 32.00 | 97.31 | 807.01 | 2839 | ! cDNA sequence encoding CEA com |

seq_name: N_Geneseq_36:Q29167

ID Q29167 standard; DNA; 270 BP.

DT 09-MAR-1993 (first entry)

KW Human leukocyte antigen; transge-
expression: ss

25-MAR-1992:

PR 03-AUG-1990; JP-207329.

DR WPI; 92-342893/42.

PT analysis of expression of gene

The sequence shows the exon 2 at the 5' end of the cDNA, which is identical to the sequence of the exon 2 of the *h* gene.

CC human mammals, pref. rat or mouse

embryo to generate transgenic no

non-human mammals contg. HLA-BW

SQ Sequence 270 BP; 59 A; 8

Quality: 49:00
Ratio: 4:900

[illegible]

US-08-653-294-8 x Q29167 ..

Align seg 1/1 to: Q29167 from: 1

I AFGGIuAshLeuArgIIeAlaLeuArgII

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1

| ID | Q01834 standard; DNA; 1086 BP. |
|-----|--------------------------------|
| 10 | Q01834 |
| 11 | Q01834 |
| 12 | Q01834 |
| 13 | Q01834 |
| 14 | Q01834 |
| 15 | Q01834 |
| 16 | Q01834 |
| 17 | Q01834 |
| 18 | Q01834 |
| 19 | Q01834 |
| 20 | Q01834 |
| 21 | Q01834 |
| 22 | Q01834 |
| 23 | Q01834 |
| 24 | Q01834 |
| 25 | Q01834 |
| 26 | Q01834 |
| 27 | Q01834 |
| 28 | Q01834 |
| 29 | Q01834 |
| 30 | Q01834 |
| 31 | Q01834 |
| 32 | Q01834 |
| 33 | Q01834 |
| 34 | Q01834 |
| 35 | Q01834 |
| 36 | Q01834 |
| 37 | Q01834 |
| 38 | Q01834 |
| 39 | Q01834 |
| 40 | Q01834 |
| 41 | Q01834 |
| 42 | Q01834 |
| 43 | Q01834 |
| 44 | Q01834 |
| 45 | Q01834 |
| 46 | Q01834 |
| 47 | Q01834 |
| 48 | Q01834 |
| 49 | Q01834 |
| 50 | Q01834 |
| 51 | Q01834 |
| 52 | Q01834 |
| 53 | Q01834 |
| 54 | Q01834 |
| 55 | Q01834 |
| 56 | Q01834 |
| 57 | Q01834 |
| 58 | Q01834 |
| 59 | Q01834 |
| 60 | Q01834 |
| 61 | Q01834 |
| 62 | Q01834 |
| 63 | Q01834 |
| 64 | Q01834 |
| 65 | Q01834 |
| 66 | Q01834 |
| 67 | Q01834 |
| 68 | Q01834 |
| 69 | Q01834 |
| 70 | Q01834 |
| 71 | Q01834 |
| 72 | Q01834 |
| 73 | Q01834 |
| 74 | Q01834 |
| 75 | Q01834 |
| 76 | Q01834 |
| 77 | Q01834 |
| 78 | Q01834 |
| 79 | Q01834 |
| 80 | Q01834 |
| 81 | Q01834 |
| 82 | Q01834 |
| 83 | Q01834 |
| 84 | Q01834 |
| 85 | Q01834 |
| 86 | Q01834 |
| 87 | Q01834 |
| 88 | Q01834 |
| 89 | Q01834 |
| 90 | Q01834 |
| 91 | Q01834 |
| 92 | Q01834 |
| 93 | Q01834 |
| 94 | Q01834 |
| 95 | Q01834 |
| 96 | Q01834 |
| 97 | Q01834 |
| 98 | Q01834 |
| 99 | Q01834 |
| 100 | Q01834 |

DT 19-MAR-1991 (first entry)
 DF sequence encoding HTA-R51 antigen

OS Homo sapiens.

PD 14-FEB-1990.

PR 11-AUG-1988; JP-200758.

FI KNO K, TAXIGUCHI,
DR WPT: 90-046289/07.

transformed cells, useful for DNase

CC The HLA class I DNA can be used.

CC obtained by introducing these DN
50 1085 BB. 324 1.

alignment_scores:
 Quality: 49.00 Length: 10
 Ratio: 4.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-8 x Q01834 ..

Align seg 1/1 to: Q01834 from: 1 to: 1086

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10

|||||

294 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 323

seq_name: N_Geneseq_36:Q01822

seq_documentation_block:

ID Q01822 standard; DNA; 1086 BP.

AC Q01822;

DT 19-MAY-1991 (first entry)

DE Sequence encoding HLA-B*52 antigen.

KW Probe; HLA class I DNA; immunogen; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT cds 1..1086

FT /*tag= a

FN EP-354580-A.

PD 14-FEB-1990.

PF 10-AUG-1989.

PR 11-AUG-1988; JP-200758.

PA (OLYU) Olympus Optical Co., Ltd.

PI Kano K, Takiguchi;

DR WPI; 90-046289/07.

DR P-PSDB; R03142.

PT New DNA for class I human leucocyte antigens and derived probes and transformed cells, useful for DNA typing, as immunogens etc.

PS Claim 2; pp11-12; 23pp; English.

CC The HLA class I DNA can be used as a source of probes for use in DNA typing. Transformed cells, which are useful as immunogens, can be obtained by introducing these DNAs into eucaryotic cells.

SQ Sequence 1086 BP; 223 A; 335 C; 358 G; 170 T;

alignment_scores:

Quality: 49.00 Length: 10
 Ratio: 4.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-8 x Q01822 ..

Align seg 1/1 to: Q01822 from: 1 to: 1086

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10

|||||

294 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 323

seq_name: N_Geneseq_36:Q05693

seq_documentation_block:

ID Q05693 standard; DNA; 1089 BP.

AC Q05693;

DT 03-JAN-1991 (first entry)

DE HLA-B*51 gene for production of monoclonal antibodies.

KW Allotype specific monoclonal anti-HLA antibodies; hybridomas;

KW transgenic animals; HLA-B*51 gene; ss.

FH Key Location/Qualifiers

FT exon 1..73

FT /*tag= a

FT /number=1

FT 74..343

FT exon

FT /*tag= b
 FT /number=2
 FT /note="alpha 1-domain"
 FT 344..619
 FT /*tag= c
 FT /number=3
 FT /note="alpha 2-domain"
 FT 620..895
 FT /*tag= d
 FT /number=4
 FT /note="alpha 3-domain"
 FT 896..1012
 FT /*tag= e
 FT /number=5
 FT 1013..1042
 FT /*tag= f
 FT /number=6
 FT 1043..1089
 FT /*tag= g
 FT /number=7

FN EP-383183-A.

PD 22-AUG-1990.

PF 07-FEB-1990; 102424.

PR 08-FEB-1989; JP-Q29313.

PA (OLYU) OLYMPUS OPTICAL KK.

PI Takiguchi M;

DR WPI; 90-255479/34.

PT Allotype specific monoclonal anti- HLA antibodies prodn. - using hybridomas derived from transgenic animals carrying HLA gene and immunised with HLA antigen of different allotype

PS Disclosure; Fig 1 A-G; 20pp; English.

CC The human HLA-B*51 gene was injected into fertilised mouse eggs and then these introduced into the uterus of a pseudo pregnant mouse.

CC The young were tested to ensure incorporation of the gene into the chromosome, and one of them mated 3 times with a normal male to

CC produce 16 young, seven of which carried the HLA-B*51 gene.

CC The transgenic offspring were immunised with HLA antigen.

CC The spleen lymphocytes were fused with myeloma cells. Hybridomas

CC producing antibodies were selected.

CC See also Q05701.

SQ Sequence 1089 BP; 224 A; 335 C; 357 G; 173 T;

alignment_scores:

Quality: 49.00 Length: 10
 Ratio: 4.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-8 x Q05693 ..

Align seg 1/1 to: Q05693 from: 1 to: 1089

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10

|||||

295 CGAGAGAACCTGCGGATCGCGCTCCGCTAC 324

seq_name: N_Geneseq_36:Q05701

seq_documentation_block:

ID Q05701 standard; DNA; 1089 BP.

AC Q05701;

DT 03-JAN-1991 (first entry)

DE HLA-B*52 gene for production of monoclonal antibodies.

KW Allotype specific monoclonal anti-HLA antibodies; hybridomas;

KW transgenic animals; HLA-B*52 gene; ss.

FH Key Location/Qualifiers

FT exon 1..73

FT /*tag= a

FT /number=1

FT 74..343

FT /*tag= b

FT /number=2

```

FT exon /note="alpha 1-domain"
FT 344..619
FT /tag= c
FT /number=3
FT /note="alpha 2-domain"
FT 620..895
FT /tag= d
FT /number=4
FT /note="alpha 3-domain"
FT 896..1012
FT /tag= e
FT /number=5
FT 1013..1042
FT /tag= f
FT /number=6
FT 1043..1089
FT /tag= g
FT /number=7
FT
PN EP-383183-A.
PD 22-AUG-1990.
PF 07-FEB-1990: 102424.
PR 08-FEB-1989: JP-029313.
PA (OLYU ) OLYMPUS OPTICAL KK.
PI Takiguchi M;
DR WPI; 90-255479/34.
PT Allotype specific monoclonal anti- HLA antibodies prodn. - using
PT hybridomas derived from transgenic animals carrying HLA gene and
PT immunised with HLA antigen of different allotype
PS Disclosure: Fig 1 A-G; 20pp; English.
CC The human HLA-Bw52 gene was introduced into mouse L cells and
CC then these cells used to immunise one of the transgenic mice
CC (See Q05693).
CC The spleen lymphocytes were fused with myeloma cells (P3x63-Ag8.653).
CC Hybridomas producing antibodies were selected.
SQ Sequence 1089 BP; 223 A; 336 C; 359 G; 171 T;

alignment_scores:
Quality: 49.00 Length: 10
Ratio: 4.900 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-8 x Q05701 ..
Align seg 1/1 to: Q05701 from: 1 to: 1089

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||
295 CGAGAGAACCTGCGGATCGGCTCCGCTAC 324

seq_name: N_Geneseq_36:Q12114

seq_documentation_block:
ID Q12114 standard; DNA; 1089 BP.
AC Q12114.
DT 29-AUG-1991 (first entry)
DE HLA-Bw53 exon.
KW Human leukocyte antigen; probe; major histocompatibility complex;
KW MHC; class I; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..1089
FT /tag= a
PN J03112487-A.
PD 14-MAY-1991.
PF 22-SEP-1989: 247697.
PR 22-SEP-1989: JP-247697.
PA (OLYU ) OLYMPUS OPTICAL KK.
DR WPI; 91-182991/25.
PT HLA-Bw53 gene.
PT immunisation, identifying specificity of antiserum etc.

```

```

PS Claim 1; Page 1; 11pp; Japanese.
CC Probes comprising part of the sequence can be used to identify
CC Class I genes. The DNA can be expressed for immunisation of
CC animals and prodn. of monoclonal antibodies specific for the
CC HLA-Bw53 antigen. See also J03112485 and J03112486.
SQ Sequence 1089 BP; 222 A; 337 C; 356 G; 174 T;

alignment_scores:
Quality: 49.00 Length: 10
Ratio: 4.900 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-653-294-8 x Q12114 ..
Align seg 1/1 to: Q12114 from: 1 to: 1089

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||
295 CGAGAGAACCTGCGGATCGGCTCCGCTAC 324

seq_name: N_Geneseq_36:X33945

seq_documentation_block:
ID X33945 standard; DNA; 2056 BP.
AC X33945.
DT 30-JUN-1999 (first entry)
DE Human HCMV inducible gene, SEQ ID NO 17.
KW HCMV inducible gene; cig; human; human cytomegalovirus; interferon;
KW anti-viral therapy; anti-HCMV therapy; detection; diagnosis;
KW drug screening; ds.
OS Homo sapiens.
PN WO9913075-A2.
PD 18-MAR-1999.
PF 08-SEP-1998; U18638.
PR 22-SEP-1997; US-059725.
PR 08-SEP-1997; US-058180.
PA (UYPH-) UNIV PRINCETON.
PI Cong J, Schenk T, Zhu H;
DR WPI; 99-243729/20.
DR P-PSDB; Y05375.
PT New isolated human genes
PS Claim 2: Page 126-129; 184pp; English.
CC This sequence represents a human gene of the invention, that is induced
CC to express by both HCMV and interferon (IFN), designated HCMV-inducible
CC genes (cig or cigs). The invention also relates to genes that are
CC repressed in the presence of HCMV infection, designated HCMV-repressible
CC genes (crg or crgs). The products can be used to obtain agents which can
CC be used for anti-viral therapy, particularly anti-HCMV therapy. They can
CC also be used for the development of drugs that would allow for higher
CC dosage IFN treatments without the concomitant toxicity normally
CC associated with administering high levels of IFN. The products can also
CC be used for detection, diagnosis and drug screening.
SQ Sequence 2056 BP; 645 A; 437 C; 526 G; 448 T;

alignment_scores:
Quality: 35.00 Length: 10
Ratio: 3.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 70.000

alignment_block:
US-08-653-294-8 x X33945 ..
Align seg 1/1 to: X33945 from: 1 to: 2056

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||
1665 AGGGAAGCTTTCATGCTTCCTAAGGTAC 1694

seq_name: N_Geneseq_36:N70935

```

```

seq_documentation_block:
ID N70935 standard; DNA; 1026 BP.
AC N70935;
DT 10-APR-1991 (first entry)
DE Sequence encoding the human histocompatibility antigen HLA B27.
KW Rheumatic disorder; genetic screening; diagnosis;
KW ankylosing spondylitis; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..1026
FT CD3
FT DE3542024-A.
FT PD 04-JUN-1987.
FT PF 28-NOV-1985; 542024.
FT PR 28-NOV-1985; DE-542024.
FT PR 21-DEC-1985; DE-545576.
FT PA (BEHW ) BEHRINGWERKE AG.
FT PI Riethmuller G, Meo T, Weiss E, Szots H;
DR WPI; 87-157893/23.
DR P-PSDB: P70590.
PT DNA coding for antigen HLA B27 - and diagnostic reagents contg.
PT such DNA, antigen or antibody
PS Claim 2; Page 4; 5pp; German.
CC The DNA may be used as a hybridisation probe for detecting the HLA
CC B27 gene, e.g. for assessing susceptibility to rheumatic disorders
CC such as ankylosis spondylitis, or may be used to transform cells
CC for prodn. of HLA B27. The HLA B27 may be used to detect HLA B27
CC antibody in human serum, or to produce mono- or polyclonal HLA B27
CC antibodies for use in immunoassay.
SQ Sequence 1026 BP; 213 A; 307 C; 344 G; 162 T;

alignment_scores:
Quality: 34.00 Length: 10
Ratio: 4.250 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 70.000

alignment_block:
US-08-653-294-8 x N70935 ..
Align seg 1/1 to: N70935 from: 1 to: 1026
1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||:||||| |||||||
223 CGAGAGGACCTGCGGACCTGCTCGCTAC 252

seq_name: N_Geneseq_36:N70225

seq_documentation_block:
ID N70225 standard; DNA; 3874 BP.
AC N70225;
DT 03-APR-1991 (first entry)
DE Sequence of genomic DNA encoding human histocompatibility antigen
DE HLA-B 27.
KW Ankylosing spondylitis; rheumatic disorder; diagnosis; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Intron 518..590
FT Intron /*tag= a
FT Intron 720..989
FT Intron /*tag= b
FT Intron 1090..1506
FT Intron /*tag= c
FT Intron 1922..2357
FT Intron /*tag= d
FT Intron 2450..2566
FT Intron /*tag= e
FT Intron 3009..3041
FT Intron /*tag= f
FT Intron 3148..3191
FT Intron /*tag= g
FT EP-226069-A.
FT PN 24-JUN-1987.
FT PF 21-NOV-1986; 116139.

```

```

PR 01-JAN-1985; DE-542024.
PR 21-DEC-1985; DE-545576.
PR (BEHW ) BEHRINGWERKE AG.
PI Szots H, Weiss E, Dörner C, Lang M, Meo T, Riethmuller G;
DR WPI; 87-171469/25.
DR P-PSDB; P70155.
PT DNA coding for human histocompatibility antigen HLA-B 27 - useful
PT for diagnosis and antigen and antibody prodn.
PS Claim 1; P6; 13pp; German.
CC The DNA may be used to detect the HLA-B 27 gene (opt. mutated) in
CC human genetic material. The HLA-B 27 may be used to detect anti-HLA-
CC B 27 antibodies in human serum. The antibodies may be used to
CC determine HLA-B 27 levels in human serum, e.g. for diagnosis of
CC rheumatic disorders, esp. ankylosing spondylitis.
CC Sequence 3874 BP; 751 A; 1094 C; 1171 G; 858 T;

alignment_scores:
Quality: 34.00 Length: 10
Ratio: 4.250 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 70.000

alignment_block:
US-08-653-294-8 x N70225 ..
Align seg 1/1 to: N70225 from: 1 to: 3874
1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||:||||| |||||||
941 CGAGAGGACCTGCGGACCTGCTCGCTAC 970

seq_name: N_Geneseq_36:T61639

seq_documentation_block:
ID T61639 standard; DNA; 6553 BP.
AC T61639;
DT 05-JUN-1997 (first entry)
DE HLA B27 consensus sequence.
DE HLA B27; seronegative spondylarthropathy; ankylosing spondylitis;
KW Reiter's syndrome; arthritis; acute anterior uveitis; diagnosis;
KW ss; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Key 3968..6653
FT mrna /*tag= a
FT mrna /*note= "HLA-B27 3' flanking region, downstream of
FT 3' untranslated region"
FT mrna 4112..4556
FT mrna /*tag= b
FT mrna /*note= "3' flanking region diagnostic for genetic
FT predisposition to SNSA"
FT mrna 4270..4556
FT mrna /*tag= b
FT mrna /*note= "3' flanking region diagnostic for genetic
FT predisposition to SNSA"
FT misc_difference 4495
FT mrna /*tag= d
FT mrna /*note= "absence of cytosine at this site is
FT indicative of a predisposition to SNSA"

WO9709450-A1.
PD 13-MAR-1997.
PD 16-AUG-1996; U13256.
PR 01-SEP-1995; US-522942.
PA (CEDA-) CEDARS SINAI MEDICAL CENT.
PI Tyan DB;
DR WPI; 97-192924/17.
PT Detecting pre-disposition to seronegative spondylarthropathies -
PT from the absence of a C residue at a specific position in the
PT 3'-flanking region of the HLA B27 allele
PS Claim 1; Page 52-56; 68pp; English.
CC Genetic predisposition to seronegative spondylarthropathies (SNSA)
CC is detected by determining the absence of a cytosine nucleotide in
CC the 3' flanking region (see also T61647-48) of an HLA-B gene at a

```

CC position corresponding to nucleotide 4495 of the HLA-B27 consensus
 CC sequence given in T61639. Probes and primers (see also T61640-46)
 CC based on this region can be used in diagnostic assays to detect the
 CC genetic predisposition to SNSA, and permit the distinction of B27+
 CC individuals who are resistant to SNSA from B27+ normal individuals
 CC who are susceptible (but as yet unaffected) to such diseases.
 SQ Sequence 6553 BP; 1443 A; 1619 C; 2017 G; 1474 T;

alignment_scores:
 Quality: 34.00 Length: 10
 Ratio: 4.250 Gaps: 0
 Percent Similarity: 80.000 Percent Identity: 70.000

alignment_block:

US-08-653-294-8 x T61639 ..

Align seg 1/1 to: T61639 from: 1 to: 6553

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10

|||||:|||||
 1102 CGAGAGACCTCGGAGCCCTGCTCCGCTAC 1131

seq_name: N_Geneseq_36.V73804

seq_documentation_block:

ID V73804 standard; DNA; 35100 BP.

AC V73804;
 DT 25-FEB-1999 (first entry)
 DE KSHV LUR DNA (nucleotides 70,201-105,300).
 KW Kaposi's sarcoma; acquired immune deficiency syndrome; AIDS; DHFR; Bcl-2;
 KW dihydrofolate reductase; LUR: long unique region; vaccine; prophylaxis;
 KW diagnosis; treatment; HHV8; transactivator; LCTP; dUTPase; Isp-II;
 KW DNA replication protein I; immediate early protein II; vRFL1; ICSBP;
 KW phosphoprotein; DNA replication protein II; ribonucleotide reductase;
 KW assembly/DNA maturation protein; tegument protein; ds.
 OS Kaposi's sarcoma-associated herpesvirus.
 PN US5849564-A.
 PD 15-DEC-1998.
 PF 29-NOV-1996.
 PR 29-NOV-1996; US-770379.

PI (UYCO) UNIV COLOMBIA NEW YORK.

PT Kaposi's sarcoma-associated herpes virus nucleic acid - encodes

PT or diagnosis of Kaposi's sarcoma

PT dihydrofolate reductase and is useful for treatment, prophylaxis

PT disclosure; Column 125-156; 109pp; English.

PS This sequence is a fragment of the Kaposi's sarcoma-associated

PS herpesvirus (KSHV) LUR (long unique region). This fragment contains

PS coding regions for ORF48, ORF49, ORF50 which encodes a transactivator

PS (LCTP), K8, ORF52, ORF53, ORF54 which encodes dUTPase, ORF55, ORF56

PS which encodes DNA replication protein I, ORF57 which encodes immediate

PS early protein II (IEP-II), K9 which encodes vRFL1 (ICSBP), K10, K11,

PS ORF58 which encodes a phosphoprotein, ORF59 which encodes DNA replication

PS protein II, ORF60 which encodes a small ribonucleotide reductase, ORF61

PS which encodes a large ribonucleotide reductase, ORF62 which encodes an

PS assembly/DNA maturation protein, ORF63 which encodes tegument protein

PS II, ORF64 which encodes tegument protein III. KSHV is a new human

PS Herpesvirus (HHV8) believed to cause Kaposi's sarcoma (KS) which is the

PS most common form of neoplasm occurring in persons with acquired immune

PS deficiency syndrome (AIDS). The DHFR protein is useful for vaccination,

PS prophylaxis, diagnosis and treatment of a subject with Kaposi's sarcoma

PS and for detecting expression of a DNA virus associated with Kaposi's

PS sarcoma in a cell.

SQ Sequence 35100 BP; 8632 A; 10010 C; 8650 G; 7808 T;

alignment_scores:

Quality: 34.00 Length: 9

Ratio: 4.250 Gaps: 0

Percent Similarity: 88.889 Percent Identity: 77.778

alignment_block:

US-08-653-294-8 x V73804/rev ..

Align seg 1/1 to reverse of: V73804 from: 1 to: 35100

1 ArgGluAsnLeuArgIleAlaLeuArg 9

|||||:|||||
 3575 CGGAGAAATCTCCGGTGTCTCTGCGA 3549

seq_name: N_Geneseq_36.V19941

seq_documentation_block:

ID V19941 standard; DNA; 137507 BP.

AC V19941;
 DT 03-AUG-1998 (first entry)
 DE KSHV long unique coding region and terminal repeat.
 KW KSHV; HHV8; human herpes virus 8; macrophage inflammatory protein II;
 KW interleukin-6; IL-6; interferon regulatory factor; rheumatoid arthritis;
 KW complement-binding protein; glycoprotein; capsid protein IV; infection;
 KW immediate early protein; Kaposi's sarcoma; protective vaccine; lymphoma;
 KW lymphoproliferative disease; leukemia; splenomegaly; mycosis fungoides;
 KW HIV immune status; anti-inflammatory agent; therapy; ds.
 OS Kaposi's sarcoma-associated herpes virus.
 FH Key Location/Qualifiers
 CDS 1142..2794
 FT /tag= a
 FT /product= complement-binding protein
 FT 8699..11236
 FT /tag= b
 FT /product= glycoprotein B
 FT complement (17261..17875)
 FT /tag= c
 FT /product= interleukin 6
 FT complement (21548..21832)
 FT /tag= d
 FT /product= macrophage inflammatory protein II
 FT complement (27137..27424)
 FT /tag= e
 FT /product= interferon regulatory factor 1
 FT 28661..29741
 FT /tag= f
 FT /product= protein T1.1
 FT complement (58976..60175)
 FT /tag= g
 FT /product= glycoprotein M
 FT complement (69412..69915)
 FT /tag= h
 FT /product= glycoprotein L
 FT complement (88410..88910)
 FT /tag= i
 FT /product= interferon regulatory factor 2
 FT 89600..90541
 FT /tag= j
 FT /product= interferon regulatory factor 3
 FT 90173..90643
 FT /tag= k
 FT /product= glycoprotein X
 FT complement (93636..94127)
 FT /tag= l
 FT /product= interferon regulatory factor 4
 FT complement (111931..112443)
 FT /tag= m
 FT /product= capsid protein IV
 FT complement (123808..127296)
 FT /tag= n
 FT /product= immediate early protein

WO9804576-A1.

PD 05-FEB-1998.

PF 22-JUL-1997; U13346.

PR 29-NOV-1996; US-757669.

PR 25-JUL-1996; US-686243.

PR 25-JUL-1996; US-686349.

PR 25-JUL-1996; US-686350.

PS Claim 1: Page 170: 215pp: English.
 CC X51701-55 encode human secreted proteins. The polynucleotides and
 CC their corresponding secreted polypeptides are useful for preventing,
 CC treating or ameliorating medical conditions, e.g. by protein or gene
 CC therapy. Pathological conditions can also be diagnosed by determining
 CC the amount of the new polypeptides in a sample or by determining the
 CC presence of mutations in the new polynucleotides. Specific uses are
 CC described for each polynucleotide, based on which tissues they are
 CC most highly expressed in, and include developing products for the
 CC diagnosis or treatment of cancer, immune disorders, infection,
 CC inflammatory disorders, skin disorders, tumours, atherosclerosis,
 CC restenosis, autoimmune disorders, Alzheimer's disease, peripheral
 CC neuropathies, trauma, spinal cord injuries, allergy, hematopoietic
 CC disorders, skeletal disorders, neurological disorders, arthritic
 CC disorders, asthma, immunodeficiency diseases, AIDS and transplant
 CC rejection. The polypeptides are also useful for identifying their
 CC binding partners.
 SQ Sequence 978 BP; 285 A; 256 C; 197 G; 240 T;

alignment_scores:
 Quality: 33.00 Length: 10
 Ratio: 3.667 Gaps: 0
 Percent Similarity: 90.000 Percent Identity: 70.000

alignment_block:
 US-08-653-294-8 x X51732/rev ..

Align seg 1/1 to reverse of: X51732 from: 1 to: 978

```

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||:|||||:|||||:|||||:
99 AGGAGGAGCTGAGATCCAACTCGGTGG 70

```

seq_name: N_Geneseq_36:N80055

seq_documentation_block:
 ID N80055 standard; DNA; 1011 BP.
 AC N80055;
 DT 10-OCT-1990 (first entry)
 DE Beta-glucuronidase-contg insert of E.coli plasmid pBG1
 KW E.coli MS371 beta-glucuronidase (BG);
 KW prokaryotic expression system: ss.
 OS Escherichia coli.
 FH Key Location/Qualifiers
 FT cds 502..1011
 FT /*tag= a
 FT rbs 485..495
 FT /*tag= b
 FT /label= strong RBS
 FT /note="start and end not defined"
 FT misc_difference 411..436
 FT /*tag= c
 FT /rpt_type=imperfect 26bp dyad
 FT /note="possible repressor binding function"
 FT misc_rna 231..552
 FT /*tag= d
 FT /note="confers high expression activity"

PN US4721671-A.
 PD 26-JAN-1988.
 PF 26-DEC-1984; 686344.
 PR 26-DEC-1984; US-686344.
 PA (REPL-) Repligen Corp.
 PI Anillonis A, Palmer JL;
 DR WPI; 88-049643/07.
 DR P-PSDB: P80057.
 PT Efficient prokaryotic expression system -
 PT using DNA coding for part or all of the amino acids of
 PT Escherichia coli beta-glucuronidase
 PS Claim 1: Page 12; 11pp: English.
 CC E.coli transformed with expression vectors contg this sequence
 CC produce BG at a high level (BG= 50% of total cellular protein).
 CC The efficient BG promoter sequence can be placed upstream of DNA

CC coding for other proteins eginsulin. Fusion proteins are then
 CC produced at high levels.
 CC The region from 231-552 or any part of it can be used in an
 CC expression system to enhance prodn of proteins.
 SQ Sequence 1011 BP; 245 A; 220 C; 253 G; 293 T;
 alignment_scores:
 Quality: 33.00 Length: 10
 Ratio: 3.667 Gaps: 0
 Percent Similarity: 90.000 Percent Identity: 60.000
 alignment_block:
 US-08-653-294-8 x N80055/rev ..
 Align seg 1/1 to reverse of: N80055 from: 1 to: 1011

```

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||:|||||:|||||:|||||:
121 AGAGAGAGATCAGGATCGGTTAAGATAC 92

```

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-8 to: EST:* out_format : pfs

Date: Feb 8, 2000 4:02 AM

About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=frame.p2n.model -DEV=xlp
-O/cn1/USF0.spool/US08653294/runat_04022000_160700_15770/app_query.fasta.1
-DB-EST -QWMT-fastap -SUFFIX=rs -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000
-GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MAPEXT=blosum62 -TRANS-human40.cdi
-LIST=45 -DOALIGN=200 -THRSCORE=pct -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294
-NCPU=3 -ICPU=3 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-8

Query length: 10

Database: EST*

Database sequences: 4538634

Database length: 1887831982

Search time (sec): 8553.360000

score_list:

| Sequence | Strd | Orig | ZScore | Escore | Len | Documentation |
|--------------------|------|-------|--------|--------|-----|----------------------------------|
| gb_est8:C03945 | + | 49.00 | 190.73 | 0.0739 | 232 | C03945 C03945 Human heart cDNA |
| gb_est10:A151891 | + | 49.00 | 189.90 | 0.0822 | 255 | A151891 Z00106.r1 Stratagene |
| gb_est11:AA263158 | + | 49.00 | 188.99 | 0.0923 | 283 | AA263158 PM0534 KGI-a Lambda 2 |
| gb_est6:D82221 | - | 49.00 | 186.54 | 0.1265 | 375 | D82221 HUMHBC4626 Human pancrea |
| gb_est26:A1359260 | - | 49.00 | 182.18 | 0.2213 | 618 | A1359260 QV27B07.x1 NCI_CGAP.Bi |
| gb_est31:A1696864 | - | 49.00 | 180.51 | 0.2740 | 748 | A1696864 w674h11.x1 NCI_CGAP.Pa |
| gb_est10:A147151 | + | 45.00 | 166.52 | 1.65 | 581 | AA147151 w32806.r1 Stratagene |
| gb_est23:A124815 | + | 44.00 | 163.42 | 1.90 | 414 | A124815 am5606.x1 Johnston fi |
| gb_est6:D82189 | + | 43.00 | 161.35 | 3.20 | 415 | D82189 HUMHBC4524 Human pancrea |
| gb_est2:R13904 | + | 42.00 | 156.42 | 6.02 | 459 | R13904 yf62c03.r1 Soares infant |
| gb_gss9:AQ164696 | - | 41.00 | 151.25 | 11.68 | 522 | AQ164696 HS_3006.B2.D07.T7 CIT |
| gb_est22:A1007865 | + | 40.00 | 151.34 | 11.55 | 325 | A1007865 EST202316 Normalized r |
| gb_est1:T24384 | - | 40.00 | 150.66 | 12.59 | 351 | T24384 crs1519 lambdaZAPST Ric |
| gb_est26:AA817822 | - | 40.00 | 146.90 | 20.40 | 540 | AA817822 UI-R-A0-a8-07-0-UI.S |
| gb_gss4:AA0570343 | - | 39.00 | 142.51 | 35.86 | 562 | AA0570343 HS_5352.B1.G08.T7A RPO |
| gb_est37:A1999216 | - | 39.00 | 141.68 | 39.88 | 618 | A1999216 701534929 A. thaliana, |
| gb_gss7:AA096303 | - | 39.00 | 139.33 | 53.91 | 809 | AA096303 HS_3072.B2.G05.T7C CIT |
| gb_gss14:AA0573870 | + | 39.00 | 138.83 | 57.42 | 856 | AQ573870 nbxb0085C03f CUGI Rice |
| gb_gss13:AA0459514 | + | 38.00 | 138.95 | 56.57 | 531 | AQ459514 HS_5085.A1.B10.T7A RPO |
| gb_est6:D82177 | + | 37.00 | 139.05 | 55.84 | 330 | D82177 HUMHBC4504 Human pancrea |
| gb_gss1:CN0005YP | + | 37.00 | 138.39 | 60.78 | 356 | AL089023 Arabidopsis thaliana g |
| gb_est27:A1466429 | + | 37.00 | 138.15 | 62.70 | 366 | A1466429 vx35804.y1 Stratagene |
| gb_gss3:R87367 | + | 37.00 | 135.24 | 91.09 | 511 | B77367 T305TFB TAMU Arabidopsi |
| gb_gss13:AA0647244 | + | 37.00 | 134.74 | 97.10 | 541 | AA0647244 HS_5210.A2.B02.T7A RPO |
| gb_est20:AA881004 | - | 37.00 | 134.22 | 103.75 | 574 | AA881004 vx35804.r1 Stratagene |
| gb_est9:AA084425 | - | 36.00 | 143.47 | 31.67 | 125 | AA084425 zf76a06.r1 Soares pine |
| gb_gss9:AA0163964 | + | 36.00 | 134.46 | 100.58 | 351 | AQ163964 HS_2270.B1.A09.MF CIT |
| gb_est8:AA008970 | + | 36.00 | 134.41 | 101.22 | 353 | AA008970 mh03d10.r1 Soares mous |
| gb_gss3:BA1200 | + | 36.00 | 133.57 | 112.84 | 389 | BA1200 HS-1053-A2-E02-MF.abi CI |
| gb_gss15:AA0592066 | + | 36.00 | 131.02 | 156.49 | 521 | AQ640210 927P1-2F2.TP 927P1 Try |
| gb_gss15:AA0592066 | + | 36.00 | 130.97 | 157.50 | 524 | AQ592066 HS_5363.A1.G04.T7A RPO |
| gb_gss14:AA0577992 | - | 36.00 | 128.23 | 223.72 | 717 | AQ577992 nbxb0092C23r CUGI Rice |
| gb_est11:AA432468 | - | 35.00 | 142.59 | 35.49 | 87 | AA432468 ve88f09.s1 Knowles SPC |
| gb_gss15:AA0640210 | + | 35.00 | 139.85 | 50.39 | 119 | AQ538221 RPCI-11-356J14.TJ RPCI |
| gb_gss14:AA0538221 | + | 35.00 | 132.51 | 129.19 | 276 | AV113853 AV113853 Mus musculus |
| gb_est33:AV113853 | + | 35.00 | 131.81 | 141.30 | 299 | AV104196 AV104196 Mus musculus |
| gb_est11:T57405 | - | 35.00 | 131.64 | 144.48 | 305 | T57405 yb51c10.r1 Stratagene fe |
| gb_est29:AA1595204 | + | 35.00 | 130.90 | 158.87 | 322 | A1595204 mk18408.y1 Soares mous |
| gb_est20:AA837121 | + | 35.00 | 130.41 | 169.08 | 351 | AA837121 ocl905.s1 NCI_CGAP.GO |
| gb_est13:AA332511 | + | 35.00 | 130.19 | 173.94 | 360 | AA332511 EST36483 Embryo, 8 wee |
| gb_est5:HA9302 | + | 35.00 | 130.12 | 175.56 | 363 | H89302 yw25c11.r1 Morton Fetal |
| gb_gss6:AA0833080 | - | 35.00 | 129.49 | 190.24 | 390 | AQ833080 HS_5499.A2.H02.SP6 RH |

gb_est3:R33700 + 35.00 128.97 203.39 414 R33700 yh78d08.r1 Soares pla
gb_est19:AA786878 - 35.00 128.89 205.59 418 AA786878 mh06a1.r1 Aspergil
gb_gss10:AQ188490 - 35.00 128.89 205.59 418 AQ188490 HS_3228.B2.G04_MR C

seq_name: gb_est8:C03945

seq_documentation_block: 232 bp mRNA EST 30-JUL-1996
LOCUS C03945 Human heart cDNA (Ynakamura) Homo sapiens cDNA clone
DEFINITION 3NHC2454, mRNA sequence.

ACCESSION C03945
VERSION C03945.1 GI:1467196
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 232)
AUTHORS Tanaka,T., Ogiwara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and Nakamura,Y.

TITLE Construction of a normalized directionally cloned cDNA library from adult heart and analysis of 3040 clones by partial sequencing
JOURNAL Genomics 35 (1), 231-235 (1996)
MEDLINE 96299762
COMMENT On Oct 24, 1995 this sequence version replaced gi:1040105.

Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.

FEATURES

source
1..232
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="3NHC2454"
/clone_lib="Human heart cDNA (Ynakamura)"
/dev_stage="adult"
/note="Organ heart; normalized directionally cloned cDNA from adult heart"

BASE COUNT 55 a 77 c 68 g 32 t
ORIGIN

alignment_scores:

Quality: 49.00 Length: 10
Ratio: 4.900 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-8 x C03945
Align seg 1/1 to: C03945 from: 1 to: 232
1 ArgGluAsnLeuArgileAlaLeuArgTyr 10
|||||
40 CGAGAGAACTCGGATCGCGTCCGCTAC 59

seq_name: gb_est10:AA151891

seq_documentation_block: 255 bp mRNA EST 10-DEC-1996
LOCUS AA151891 Homo sapiens colon (#937204) Homo sapiens cDNA clone
DEFINITION Z00106.r1 Stratagene IMAGE156435 5' similar to gb:MI5497_cds1 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-24(A-9) A*2401 (HUMAN);, mRNA sequence.

ACCESSION AA151891
VERSION AA151891.1 GI:1720754
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

REFERENCE 1 (bases 1 to 255)
 AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Huitman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R., and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (9), 807-828 (1996)
 MEDLINE 97044478
 COMMENT On May 8, 1995 this sequence version replaced gi:800234.
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnlni.gov) for further information.
 Trace considered overall poor quality
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers
 1..255
 /organism="Homo sapiens"
 /db_xref="GB:4590888"
 /db_xref="taxon:9606"
 /clone="IMAGE:566435"
 /clone_lib="Stratagene colon (#937204)"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="Organ: colon; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. T-84 colonic epithelial cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3'"

BASE COUNT 57 a 70 c 75 g 44 t
 ORIGIN

alignment_scores:
 Quality: 49.00 Length: 10
 Ratio: 4.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-08-653-294-8 x AA151891 ..
 Align seg 1/1 to: AA151891 from: 1 to: 255

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
 |||||
 77 CGAGAGAACCTCGGATCGGCTCCGCTAC 106

seq_name: gb_est11:AA263158

seq_documentation_block:
 LOCUS AA263158 283 bp mRNA EST 02-JUL-1998
 DEFINITION PMY0534 KGI-1-a Lambda zap Express cDNA library Homo sapiens cDNA 5', mRNA sequence.
 ACCESSION AA263158
 VERSION AA263158
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 283)
 AUTHORS Claudio, J.O., Liew, C.C., Dempsey, A.A., Cukerman, E., Stewart, A.K., Na, E., Atkins, H.I., Iscove, N.N., and Hawley, R.G.
 TITLE Identification of sequence-tagged transcripts differentially

expressed within the human hematopoietic hierarchy
 Genomics 50 (1), 44-52 (1998)
 MEDLINE 98292493
 COMMENT On May 5, 1995 this sequence version replaced gi:797810.
 Contact: Hawley RG
 Oncology Research Laboratories
 The Toronto Hospital
 CRCS-424, 67 College St., Toronto, Ontario M5G 2M1, Canada
 Tel: 416 3403834
 Fax: 416 3403453
 Email: r.hawley@utoronto.ca
 Similar to M58636 MHC class I HLA-Bw gene. Clone was randomly picked from KGIa primary library.
 Seq primer: 5' GAATTAACCTCACTAAAGG 3'
 High quality sequence stop: 283.
 Location/Qualifiers
 1..283
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="KGI-1-a Lambda zap Express cDNA library"
 /cell_type="promyeloblast"
 /cell_line="KGI-1"
 /note="Vector: Lambda Zap Express (Stratagene); Site_1: EcoRI; Site_2: XhoI; Unidirectional cloning sites: EcoRI-XhoI. mRNA was purified from KGI-1 cell line, CDNA was synthesized using an XhoI-OligodT linker primer. EcoRI adaptors were ligated, followed by digestion with XhoI for directional cloning into predigested Lambda Zap Express".

BASE COUNT 64 a 91 c 88 g 40 t
 ORIGIN

alignment_scores:
 Quality: 49.00 Length: 10
 Ratio: 4.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-08-653-294-8 x AA263158 ..
 Align seg 1/1 to: AA263158 from: 1 to: 283

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
 |||||
 120 CGAGAGAACCTCGGATCGGCTCCGCTAC 149

seq_name: gb_est6:D82221

seq_documentation_block:
 LOCUS D82221 375 bp mRNA EST 09-FEB-1996
 DEFINITION HUHBC4626 Human pancreatic islet Homo sapiens cDNA similar to HLA-B, mRNA sequence.
 ACCESSION D82221
 VERSION D82221.1 GI:1183739
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 375)
 AUTHORS Takeda, J.
 TITLE Human pancreatic islet ESTs
 JOURNAL Unpublished (1995)
 COMMENT On Apr 14, 1993 this sequence version replaced gi:785255.
 Contact: Jun Takeda
 Institute for Molecular and Cellular Regulation, Gunma University
 3-39-15 Showa-machi, Maebashi Gunma 371, Japan
 Tel: 272-20-8856
 Fax: 272-20-8896
 Email: jtakeda@sb.gunma-u.ac.jp.
 Location/Qualifiers
 1..375
 /organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone_lib="Human pancreatic islet"
/Note="Vector: Lambda ZAPII; Site_1: Eco RI; Site_2: Xho
I; mRNA was prepared from normal adult human islets. cDNA
is directionally synthesized from the Xho I in the vector
to the EcoRI site. cDNA was size fractionated to remove
sequences <1000 bp in size."
BASE COUNT      75 a 124 c 118 g 55 t      3 others
ORIGIN

```

```

alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```
alignment_block:

```

```
US-08-653-294-8 x D82221 ..
```

```
Align seg 1/1 to: D82221 from: 1 to: 375
```

```

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||
306 CGAGAGAACCTGCGGATCGCGTCCGNTAC 335

```

```
seq_name: gb_est26:A1359260
```

```

seq_documentation_block:
LOCUS      A1359260      618 bp      mRNA      EST      15-FEB-1999
DEFINITION qy27b07.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2013205 3'
similar to gb:D32129 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
AW-66(A-10) A*6601 ALPHA (HUMAN);, mRNA sequence.
ACCESSION  A1359260
VERSION    A1359260.1 GI:4110881
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 618)
AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTGA), Tumor Gene Index
JOURNAL    Unpublished (1998)
COMMENT    Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image.html

```

```

Insert Length: 691 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 458.
Location/Qualifiers
1. 618
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2013205"
/clone_lib="NCI_CGAP_Brn23"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/Note="Organ: brain; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
TGTTACCAATCTGAAGTGGAGCGCGCATATCTTTTTTTTTTTTTTTTTT

```

```
FEATURES
source
```

```

T 3'; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT      128 a 171 c 182 g 137 t
ORIGIN

```

```

alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```
alignment_block:

```

```
US-08-653-294-8 x A1359260/rev ..
```

```
Align seg 1/1 to reverse of: A1359260 from: 1 to: 618
```

```

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
|||||
319 CGAGAGAACCTGCGGATCGCGTCCGCTAC 290

```

```
seq_name: gb_est31:A1696864
```

```

seq_documentation_block:
LOCUS      A1696864      748 bp      mRNA      EST      03-JUN-1999
DEFINITION wc74h1.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2324421 3'
similar to gb:M28205 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
B-51(B-5) B*5101 ALPHA (HUMAN);, mRNA sequence.
ACCESSION  A1696864
VERSION    A1696864.1 GI:4984764
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 748)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    On Mar 16, 1998 this sequence version replaced gi:2961758.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

```

```

Seq primer: -400P from Gibco
High quality sequence stop: 424.
Location/Qualifiers
1. 748
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2324421"
/clone_lib="NCI_CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/Note="Organ: pancreas; Vector: pcwv-sport6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"
BASE COUNT      169 a 227 c 237 g 108 t      7 others
ORIGIN

```

```
FEATURES
source
```

```

alignment_scores:
  Quality: 49.00      Length: 10
  Ratio: 4.900       Gaps: 0

```

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-653-294-8 x AI696864 ..

Align seg 1/1 to: AI696864 from: 1 to: 748

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10

|||||
137 CGAGAGAACCTCGCGATCGCGCTCGCTAC 166

seq_name: gb_est10:AA147151

seq_documentation_block: 581 bp mRNA EST 05-DEC-1996
LOCUS AA147151
DEFINITION 2032406.r1 Stratagene colon (#937204) Homo sapiens CDNA clone
IMAGE:588587 5' similar to gb:M64740 HLA CLASS I HISTOCOMPATIBILITY
ANTIGEN, A-24(A-9) A*2402 ALPHA (HUMAN);, mRNA sequence.

ACCESSION AA147151
VERSION AA147151.1 GI:1716526
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 581)

AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R.,
and Marra,M.

TITLE Generation and analysis of 280,000 human expressed sequence tags

JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393699.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 272.

FEATURES

Location/Qualifiers
1..581
/organism="Homo sapiens"
/db_xref="GDB:4620889"
/db_xref="taxon:9606"
/clone="IMAGE:588587"
/clone_lib="Stratagene colon (#937204)"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: colon; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. T-84 colonic epithelial cell line. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor
sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTT 3'"

BASE COUNT 134 a 162 c 185 g 85 t 15 others
ORIGIN
alignment_scores:
Quality: 45.00 Length: 10
Ratio: 4.500 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 90.000

alignment_block:

US-08-653-294-8 x AA147151 ..

Align seg 1/1 to: AA147151 from: 1 to: 581

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10

|||||
152 CGAGAGAACCTCGCGATCGCGCTCGCTAC 181

seq_name: gb_est23:AI124815

seq_documentation_block: 414 bp mRNA EST 11-SEP-1998
LOCUS AI124815
DEFINITION am56e06.x1 Johnston frontal cortex Homo sapiens CDNA clone
IMAGE:1539586 3' similar to gb:M24038 cds1 HLA CLASS I
HISTOCOMPATIBILITY ANTIGEN, BW-44(B-12) B*4402 (HUMAN); contains
MER22.t3 TARI repetitive element ;, mRNA sequence.

ACCESSION AI124815
VERSION AI124815.1 GI:3593329
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 414)

AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.

Wasnu-NCI Human EST Project
Unpublished (1997)

JOURNAL

COMMENT

On Jan 17, 1998 this sequence version replaced gi:1899887.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40m13 fwd. ET from Amersham.

FEATURES

Location/Qualifiers
1..414
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1539586"
/clone_lib="Johnston frontal cortex"
/sex="male"
/tissue_type="pooled frontal lobe"
/dev_stage="adult"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: brain; Vector: Bluescript SK-; Site_1:
EcoRI; Stanley Neuropathology Consortium
(www.stanleylab.org) brains S-58, S-65, S-67, S-78.
Random + oligo-dt primed into EcoRI site of ZAP II Vector.
Mass excised. Avg insert length 1.9kb. Custom library
provided by Dr. Nancy Johnston [410] 614-3918,
nlj@wustl.edu.

BASE COUNT 80 a 140 c 136 g 58 t
ORIGIN

alignment_scores:

Quality: 44.00 Length: 10
Ratio: 4.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:

US-08-653-294-8 x AI124815 ...

Align seg 1/1 to: AI124815 from: 1 to: 414

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10

|||||
284 CGAGAGAACCTCGCGACCGCGCTCGCTAC 313

seq_name: gb_est6:D82189

seq_documentation_block: 415 bp mRNA EST 09-FEB-1996
LOCUS D82189
DEFINITION HUMHC4524 Human pancreatic islet Homo sapiens cDNA similar to HLA-B, mRNA sequence.

ACCESSION D82189
VERSION D82189
KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 415)

AUTHORS Takeda J.

TITLE Human pancreatic islet ESTs

JOURNAL Unpublished (1995)

COMMENT On Apr 14, 1993 this sequence version replaced gi:785206.

Contact: Jun Takeda

Institute for Molecular and Cellular Regulation, Gunma University
3-39-15 Showa-machi, Maebashi Gunma 371, Japan

Tel: 272-20-8856

Fax: 272-20-8896

Email: jtakeda@esb.gunma-u.ac.jp.

FEATURES Location/Qualifiers

source

1..415

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="Human pancreatic islet"

/note="Vector: Lambda ZAPII; Site_1: Eco RI; Site_2: Xho I; mRNA was prepared from normal adult human islets. cDNA was directionally synthesized from the Xho I in the vector to the EcoRI site. cDNA was size fractionated to remove sequences <1000 bp in size."

BASE COUNT 79 a 134 c 133 g 61 t 8 others
ORIGIN

alignment_scores:

Quality: 43.00 Length: 10
Ratio: 4.778 Gaps: 0
Percent similarity: 90.000 Percent identity: 90.000

alignment_block:

US-08-653-294-8 x D82189 ..

Align seg 1/1 to: D82189 from: 1 to: 415

1 ArgGlusnLeuArgIleAlaLeuArgTyr 10

|||||

279 CGAGAGAACCTGGGATCGCGCTCCGNTAC 308

seq_name: gb_est2:R13904

seq_documentation_block:

LOCUS R13904 459 bp mRNA EST 12-APR-1995

DEFINITION Yf62c03.r1 Soares infant brain 1N1B Homo sapiens cDNA clone IMAGE:26801.5. similar to gb:M64742.cdsl HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A*23(A-9) A*2301 (HUMAN);, mRNA sequence.

ACCESSION R13904

VERSION R13904.1 GI:766980

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 459)

AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.

TITLE

JOURNAL

COMMENT

The WashU-Merck EST Project

Unpublished (1995)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 1875

High quality sequence stops: 384 Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

Insert Length: 1875 Std Error: 0.00

Seq primer: M13RP1

High quality sequence stop: 384.

FEATURES

source

1..459

/organism="Homo sapiens"

/db_xref="GDB:399148"

/db_xref="taxon:9606"

/clone="IMAGE:26801"

/clone_lib="Soares infant brain 1N1B"

/sex="female"

/dev_stage="73 days post natal"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: whole brain; Vector: Lfamid BA; Site_1: Not I; Site_2: Hind III; 1st strand cDNA was primed with a Not I-oligo(dT) primer [5',

AACGAGAAATCGCGCGGAGGAATTTTTTTTTTTT 3'];

double-stranded cDNA was ligated to Hind III adaptors

(Pharmacia), digested with Not I and directionally cloned

into the Not I and Hind III sites of the Lfamid BA vector.

Library went through one round of normalization. Library

constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 88 a 144 c 146 g 75 t 6 others
ORIGIN

alignment_scores:

Quality: 42.00 Length: 9
Ratio: 4.667 Gaps: 0
Percent similarity: 100.000 Percent identity: 100.000

alignment_block:

US-08-653-294-8 x R13904 ..

Align seg 1/1 to: R13904 from: 1 to: 459

1 ArgGlusnLeuArgIleAlaLeuArg 9

|||||

286 CGAGAGAACCTGGGATCGCGCTCCGC 312

seq_name: gb_gss9:AQ164696

seq_documentation_block:

LOCUS AQ164696 522 bp DNA GSS 16-OCT-1998

DEFINITION HS_3006_B2.D07_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-3006 Col-14 Row-H, genomic survey

sequence.

ACCESSION AQ164696

VERSION AQ164696.1 GI:3562891

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 522)

AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and Hood, L.

Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

TITLE

JOURNAL

```

/clone.lib="Normalized rat brain, Bento Soares"
/note="Organi: brain; Vector: pT73Pac; Site_1: EcoRI;
Site_2: NotI"
BASE COUNT      107 a      61 c      76 g      81 t
ORIGIN

alignment_scores:
  Quality:      40.00      Length:      10
  Ratio:        4.000      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 70.000

alignment_block:
  US-08-653-294-8 x AI007865      ..

  Align seg 1/1 to: AI007865 from: 1 to: 325

    1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
      |||:||||:||||:||||:||||:||||:
    207 AGGAAGAATGTCAGGTGGCATTGAGATAC 236

seq_name: gb_est1:T24384

seq_documentation_block:
  LOCUS      T24384      351 bp      mRNA      EST      28-JUL-1995
  DEFINITION crs1519 lambdaZAPST Ricinus communis cDNA clone pcrs1519, mRNA
  sequence.
  ACCESSION      T24384
  VERSION      T24384
  KEYWORDS      EST.
  SOURCE      T24384.1 GI:689202
  ORGANISM      Ricinus communis
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; Core
  eudicots; Rosidae; eurosids I; Malpighiales; Euphorbiaceae;
  Ricinus.
  REFERENCE      1 (bases 1 to 351)
  AUTHORS      Vandelaar, F.J., Turner, S. and Somerville, C.
  TITLE      Expressed sequence tags from developing castor seeds
  JOURNAL      Plant Physiol. 108, 1141-1150 (1995)
  COMMENT      Contact: Somerville CR
  Carnegie Institution
  Carnegie Institution, 290 Panama St, Stanford, CA 94305
  Tel: 4153251521
  Email: crs@andrew.stanford.edu
  Seq primer: T3.

FEATURES
  source
    1..351
    /organism="Ricinus communis"
    /strain="Baker 296"
    /db_xref="taxon:3988"
    /clone="pcrs1519"
    /clone.lib="lambdaZAPST"
    /note="Vector: lambdaZAPII; Site 1: EcoRI; Site 2: XhoI;
poly(A)+ RNA was purified from developing stage III to
stage V (Greenwood & Bewley, Can. J. Bot. 60:1751-1760,
1982) endosperm plus embryo of immature castor fruits.
cDNA was synthesized and cloned into lambdaZAPII according
to the instructions of the manufacturer (Stratagene);
synthesis was primed from the poly(A) tail, and cloned
directionally into XhoI (3') and EcoRI (5') sites. In few
cases, sequence data indicated that this directionality
was reversed. Partial cDNA clones predominate."
BASE COUNT      104 a      61 c      77 g      102 t      7 others
ORIGIN

alignment_scores:
  Quality:      40.00      Length:      10
  Ratio:        4.000      Gaps:      0
  Percent Similarity: 100.000      Percent Identity: 80.000

alignment_block:

```

US-08-653-294-8 x T24384/rev ..

Align seg 1/1 to reverse of: T24384 from: 1 to: 351

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
 |||:::|||||
 303 AGACAAATTCAGGATTCAGCTAAGATAC 274

seq_name: gb_est26:AA817822

seq_documentation_block: 540 bp mRNA EST 03-JUL-1999
 LOCUS AA817822
 DEFINITION UI-R-A0-ae-d-07-0-UI.s1 UI-R-A0 Rattus norvegicus cDNA clone
 UI-R-A0-ae-d-07-0-UI 3' similar to gb|U92535|BTU92535 Bos taurus
 neuronal axonal membrane protein (NAP-22) mRNA, complete cds, mRNA
 sequence.
 ACCESSION AA817822
 VERSION AA817822.1 GI:4198348
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 REFERENCE 1 (bases 1 to 540)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT On Feb 17, 1998 this sequence version replaced gi:2887702.
 Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: mscoares@blue.weeg.uiowa.edu
 The sequence tag present in the cDNA between the NotI site and the
 oligo-dT track served to identify it as a clone from the normalized
 adult Brain library. cDNA Library Preparation: M. Fatima Bonaldo,
 Ph.D. Clone distribution: clones will be available through Research
 Genetics This clone is also available through the I.M.A.G.E.
 Consortium at LNL (info@image.lnl.gov). IMAGE ID=1767300 The
 following repetitive elements were found in this cDNA sequence:
 4-126. >(TAAA)n\$Simple_repeat
 Seq primer: M13 Forward
 POLYA-No.

FEATURES
 source
 1..540
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-A0-ae-d-07-0-UI"
 /clone_lib="UI-R-A0"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="vector: p7T3D-Pac (Pharmacia)" with a modified
 polylinker: site_1: Not I; Site_2: Eco RI; This library
 consists of a mixture of individually tagged normalized
 libraries constructed from rat placenta, adult lung,
 brain, liver, kidney, heart, spleen, ovary, and muscle.
 The tag is a string of 3-5 nucleotides present between the
 Not I site and the oligo-dT track which allows
 identification of the library of origin of a clone within
 the mixture."
 BASE COUNT 166 a 92 c 100 g 182 t
 ORIGIN

alignment_scores:
 Quality: 40.00 Length: 10
 Ratio: 4.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 70.000

alignment_block:

US-08-653-294-8 x AA817822 ..

Align seg 1/1 to: AA817822 from: 1 to: 540

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
 |||:::|||||
 340 AGGAAGATGTACGTTGCATTGAGATAC 369

seq_name: gb_gss14:AQ570343

seq_documentation_block: 562 bp DNA GSS 01-JUN-1999
 LOCUS AQ570343
 DEFINITION HS_5352_B1_G08_T7A RPCI-11 Human Male BAC Library Homo sapiens
 genomic clone Plate-928 Col-15 Row-N, genomic survey sequence.
 ACCESSION AQ570343
 VERSION AQ570343.1 GI:4963563
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 562)
 AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Kellar,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 COMMENT On Feb 19, 1999 this sequence version replaced gi:4146293.
 Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
 or from Resear h Genetics (info@resgen.com). BAC end Web Server:
 http://www.htsc.washington.edu
 Plate: 928 row: N column: 15
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 562.
 Location/Qualifiers
 source
 1..562
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate-928 Col-15 Row-N"
 /clone_lib="RPCI-11 Human Male BAC Library"
 /sex="male"
 /note="vector: pBACE3 6; Genomic sequence of BAC ends"

BASE COUNT 160 a 111 c 114 g 175 t 2 others
 ORIGIN

alignment_scores:
 Quality: 39.00 Length: 10
 Ratio: 3.900 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 70.000

alignment_block:

US-08-653-294-8 x AQ570343/rev ..

Align seg 1/1 to reverse of: AQ570343 from: 1 to: 562

1 ArgGluAsnLeuArgIleAlaLeuArgTyr 10
 |||:::|||||

378 AAAGAACACCTACGGTTAGCATTAGATAC 349

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:29:37 ; Search time 122.56 Seconds
(without alignments)
1.933 Million cell updates/sec

Title: US-08-653-294-9
Perfect score: 49
Sequence: 1 YRLAIRLNER 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 49 | 100.0 | 10 | 1 W47266 | Immunomodulatory p |
| 2 | 49 | 100.0 | 20 | 1 R92909 | HLA-B2702 CTL modu |
| 3 | 49 | 100.0 | 20 | 1 R92911 | HLA-B2702 CTL modu |
| 4 | 49 | 100.0 | 20 | 1 R92907 | HLA-B2702 CTL modu |
| 5 | 49 | 100.0 | 20 | 1 R95428 | HLA-B2702 84-75-84 |
| 6 | 49 | 100.0 | 20 | 1 R33778 | Immunomodulating d |
| 7 | 49 | 100.0 | 20 | 1 R33779 | Immunomodulating d |
| 8 | 49 | 100.0 | 20 | 1 R33792 | Peptide B2702.84-7 |
| 9 | 44 | 89.8 | 10 | 1 W47268 | Immunomodulatory p |
| 10 | 44 | 89.8 | 10 | 1 W47270 | Immunomodulatory p |
| 11 | 44 | 89.8 | 20 | 1 R92910 | HLA-B2702 CTL modu |
| 12 | 44 | 89.8 | 20 | 1 R92908 | HLA-B2702 CTL modu |
| 13 | 44 | 89.8 | 20 | 1 R95430 | HLA-B2702 84-75/77 |
| 14 | 44 | 89.8 | 20 | 1 W33791 | Peptide B2702.84-7 |
| 15 | 44 | 89.8 | 20 | 1 W33793 | Peptide B2702.84-7 |
| 16 | 39 | 79.6 | 10 | 1 W47272 | Immunomodulatory p |
| 17 | 36 | 73.5 | 485 | 1 R20796 | ERV-4 gC. Nucleic |
| 18 | 31 | 63.3 | 34 | 1 W76812 | N. gonorrhoeae pilC |
| 19 | 31 | 63.3 | 738 | 1 R69849 | Ethylene response |
| 20 | 31 | 63.3 | 738 | 1 R69852 | Ethylene response |
| 21 | 31 | 63.3 | 738 | 1 R69853 | Ethylene response |
| 22 | 31 | 63.3 | 738 | 1 W73121 | A. thaliana ethyle |
| 23 | 31 | 63.3 | 738 | 1 W73117 | A. thaliana ethyle |
| 24 | 31 | 63.3 | 738 | 1 W73118 | A. thaliana ethyle |
| 25 | 31 | 63.3 | 738 | 1 W73119 | A. thaliana ethyle |
| 26 | 31 | 63.3 | 738 | 1 W73120 | A. thaliana ethyle |
| 27 | 30 | 61.2 | 53 | 1 W19361 | Beta 7 integrin S3 |
| 28 | 30 | 61.2 | 376 | 1 R21416 | Carbonic anhydrase |
| 29 | 30 | 61.2 | 377 | 1 R21417 | Chlamydomonas carb |
| 30 | 30 | 61.2 | 408 | 1 W94245 | A. orientalis glyc |
| 31 | 30 | 61.2 | 408 | 1 W94246 | A. orientalis glyc |
| 32 | 29 | 59.2 | 6 | 1 W47263 | Immunomodulatory p |
| 33 | 29 | 59.2 | 6 | 1 W33781 | Peptide #2 used in |
| 34 | 29 | 59.2 | 12 | 1 R95429 | HLA-B2702 84-79-84 |

35 29 59.2 12 1 W33798 Peptide B2702.84-7
36 29 59.2 12 1 W33799 Immunomodulating d
37 29 59.2 149 1 W62901 Mutant of the firs
38 29 59.2 149 1 W51846 Amino acid sequenc
39 29 59.2 211 1 R77877 Human Cdn-2. New n
40 29 59.2 211 1 W03669 Bgk-2 protein. Scr
41 29 59.2 215 1 R38690 PGP. DNA having ge
42 29 59.2 239 1 W74405 S. aureus gidB pro
43 29 59.2 239 1 W74406 S. aureus gidB pro
44 29 59.2 247 1 W88359 Human lymphocyte a
45 29 59.2 338 1 W88361 Human lymphocyte a

ALIGNMENTS

RESULT 1
W47266
ID W47266 standard; peptide; 10 AA.
AC W47266;
DC 22-MAY-1998 (first entry)
DE Immunomodulatory peptide.
KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
KW transplant rejection; treatment; autoimmune disease.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 1..10 /note= "at least one of the amino acids is the D-isomer"
FT PN W09744052-A1.
FT PD 27-NOV-1997.
FT PF 23-APR-1997; U06705.
FT PR 22-MAY-1996; US-651650.
FT PA (STRD) UNIV LELAND STANFORD JUNIOR.
FT PI Clayberger C, Krensky AM;
FT DR WPI; 98-018220/02.
FT PT Novel immunomodulatory peptide-type compound - useful for inhibiting transplant rejection
PT Claim 10; Page 36; 41pp; English.
CC The present sequence is an immunomodulatory peptide, which comprises a Class I HLA-B alpha-1 domain sequence. It can be used in a pharmaceutical composition together with a subtherapeutic dose of an immunosuppressant, to extend the period of acceptance of a transplant from a major histocompatibility complex (MHC) unmatched donor, i.e. to inhibit transplant rejection. It can also be used in the treatment of autoimmune diseases.
CC Peptides using the D-form amino acids are more effective
CC immunomodulators than their diastereomers or enantiomers.
SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
| | | | | | | | | |
Db 1 YRLAIRLNER 10

RESULT 2
R92909
ID R92909 standard; peptide; 20 AA.
AC R92909;
DC 16-MAY-1996 (first entry)
DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84(T)).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW class I MHC; HLA-B2702.
OS Synthetic.
PN W09526979-A1.
PD 12-OCT-1995.

```

PF 05-APR-1995; U04349.
PR (STRD ) UNIV LELAND STANFORD JUNIOR.
PA Clayberger C, Krensky AM, Parham P;
PI WPI: 95-358582/46.
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
PS Example 15; Page 36; 80pp; English.
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of
CC class I major histocompatibility complex (MHC) antigens. This sequence
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class
CC I MHC HLA-B2702. These sequences can be used to extend the period of
CC acceptance by a recipient of a transplant from an MHC unmatched donor.
CC The peptides are administered to a patient in conjunction with a
CC subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.
SQ Sequence 20 AA;

Query Match 100.0%; Score 49; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00024; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 YRLAIRLNER 10
DB 1 YRLAIRLNER 10
|||||

RESULT 3
R92911
ID R92911 standard; peptide; 20 AA.
AC R92911;
DE HLA-B2702 CTL modulating peptide (B2702.84-75/84-75).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW class I MHC; HLA-B2702.
OS Synthetic.
PN W09526979-A1.
PD 12-OCT-1995.
PF 05-APR-1995; U04349.
PR (STRD ) UNIV LELAND STANFORD JUNIOR.
PA Clayberger C, Krensky AM, Parham P;
PI WPI: 95-358582/46.
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
PS Example 15; Page 36; 80pp; English.
CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of
CC class I major histocompatibility complex (MHC) antigens. This sequence
CC is a dimer of residues 75-84 of the alpha-1 domain of the class I MHC
CC HLA-B2702. These sequences can be used to extend the period of
CC acceptance by a recipient of a transplant from an MHC unmatched donor.
CC The peptides are administered to a patient in conjunction with a
CC subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.
SQ Sequence 100.0%; Score 49; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00024; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 YRLAIRLNER 10
DB 1 YRLAIRLNER 10
|||||

```

```

Query Match 100.0%; Score 49; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00024; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 YRLAIRLNER 10
DB 1 YRLAIRLNER 10
|||||

```

```

DB 1 YRLAIRLNER 10

RESULT 4
R92907
ID R92907 standard; peptide; 20 AA.
AC R92907;
DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW class I MHC; HLA-B2702.
OS Synthetic.
PN W09526979-A1.
PD 12-OCT-1995.
PF 05-APR-1995; U04349.
PR (STRD ) UNIV LELAND STANFORD JUNIOR.
PA Clayberger C, Krensky AM, Parham P;
PI WPI: 95-358582/46.
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
PS Example 15; Page 36; 80pp; English.
CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of
CC class I major histocompatibility complex (MHC) antigens. This sequence
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class
CC I MHC HLA-B2702. These sequences can be used to extend the period of
CC acceptance by a recipient of a transplant from an MHC unmatched donor.
CC The peptides are administered to a patient in conjunction with a
CC subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.
SQ Sequence 20 AA;

Query Match 100.0%; Score 49; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00024; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 YRLAIRLNER 10
DB 1 YRLAIRLNER 10
|||||

```

```

Query Match 100.0%; Score 49; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00024; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 YRLAIRLNER 10
DB 1 YRLAIRLNER 10
|||||

RESULT 5
R95428
ID R95428 standard; peptide; 20 AA.
AC R95428;
DE HLA-B2702 84-75-84 palindrome.
KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
KW cytotoxic; antigen presenting cell.
OS Synthetic.
PN W09513288-A1.
PD 18-MAY-1995.
PF 10-NOV-1994; U12985.
PR 10-NOV-1993; US-150493.
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM;
PI WPI: 95-194027/25.
PT Compens. comprising lymphoid surface membrane proteins - which may
PT inhibit cytolytic activity and differentiation of CTLs.
PS Example; Page 12; 29pp; English.
CC R95413, and R95415-R95431 represent palindromes and fragments of
CC human-leucocyte-associated antigens. This sequence represents the
CC HLA-B2702 84-75-84 palindrome. These sequences can be used to isolate
CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
CC protein associated with T-cell activation in mammalian T-cells, and is

```

CC also immunologically cross reactive with the heat shock protein Hsc70.
 CC p74 is found in a limited number of cell types, but is particularly
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable
 CC cell with an amphoteric detergent, and then passed through an affinity
 CC column containing a covalently bound HLA-B2702 palindromic peptide.
 CC Compositions comprising the extracellular fragment of p74 combined with
 CC HLA-B2702:60-84 (see R95416), induces calcium influx, and inhibits
 CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate
 CC compounds can be screened for their effect on the cytolytic activity of
 CC T-cells, by combining them with the extracellular portion of p74 and
 CC determining the amount of binding between the candidate compound and p74.
 CC Modulation of CTL activity can be inhibited in a cellular composition
 CC containing T-cells and antigen presenting cells (APCs), by adding to the
 CC mix the extracellular portion of p74, in an amount sufficient to compete
 CC with p74 for the binding of the p74 ligand.
 CC Sequence 20 AA;

Query Match 100.0%; Score 49; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.00024;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 DB 1 YRLAIRLNER 10
 |||||

RESULT 6
 W33778

ID W33778 standard; peptide; 20 AA.
 AC W33778;
 DT 19-JUN-1998 (first entry)
 DE Immunomodulating dimer peptide #1.
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplacental; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN WO9744351-Al.
 PD 27-NOV-1997.
 PF 22-MAY-1997; U08689.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI: 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases

PS Claim 16; Page 35; 41pp; English.
 CC This sequence represents a specifically claimed immunomodulating
 CC dimer peptide of the invention. A peptide-type compound or variant is
 CC claimed which has immunomodulating activity, including the N-terminal
 CC acylated and/or C-terminal amidated or esterified forms of up to 60
 CC amino acids, where the peptide-type compound comprises the formula: A-B,
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa
 CC represents amino acid. The sequence in the brackets may optionally be
 CC absent or truncated at any peptide type bond within the brackets. The
 CC compounds comprise amino acid sequences related to a Class I HLA-B
 CC alpha1 domain (positions 79-84). They can be used to inhibit cytotoxic
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in
 CC vitro. They can also be used in combination with antigenic peptides or
 CC proteins of interest to activate CTLs. They can also inhibit the
 CC proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.
 CC Sequence 20 AA;

Query Match 100.0%; Score 49; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.00024;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YRLAIRLNER 10
 DB 1 YRLAIRLNER 10
 |||||

RESULT 7
 W33779

ID W33779 standard; peptide; 20 AA.
 AC W33779;
 DT 19-JUN-1998 (first entry)
 DE Immunomodulating dimer peptide #2.
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplacental; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN WO9744351-Al.
 PD 27-NOV-1997.
 PF 22-MAY-1997; U08689.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI: 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases

PS Claim 16; Page 35; 41pp; English.

CC This sequence represents a specifically claimed immunomodulating
 CC dimer peptide of the invention. A peptide-type compound or variant is
 CC claimed which has immunomodulating activity, including the N-terminal
 CC acylated and/or C-terminal amidated or esterified forms of up to 60
 CC amino acids, where the peptide-type compound comprises the formula: A-B,
 CC where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or
 CC V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a
 CC hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa
 CC represents amino acid. The sequence in the brackets may optionally be
 CC absent or truncated at any peptide type bond within the brackets. The
 CC compounds comprise amino acid sequences related to a Class I HLA-B
 CC alpha1 domain (positions 79-84). They can be used to inhibit cytotoxic
 CC T-lymphocytes (CTL) from undesirably attacking cells in a host or in
 CC vitro. They can also be used in combination with antigenic peptides or
 CC proteins of interest to activate CTLs. They can also inhibit the
 CC proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.
 CC Sequence 20 AA;

Query Match 100.0%; Score 49; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.00024;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 DB 1 YRLAIRLNER 10
 |||||

RESULT 8
 W33792

ID W33792 standard; peptide; 20 AA.
 AC W33792;
 DT 19-JUN-1998 (first entry)
 DE Peptide B2702.84-75/75-84T tested for immunomodulating activity.
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplacental; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN WO9744351-Al.
 PD 27-NOV-1997.

PF 22-MAY-1997; U08689.
 PR (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI: 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Example 1; Page 19; 41pp; English.
 CC Peptides W3784-98 and W3778-9 were assayed for their immunomodulating
 CC activity. A peptide-type compound or variant is claimed which has
 CC immunomodulating activity, including the N-terminal acylated and/or
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where
 CC the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 SQ Sequence 20 AA;

Query Match 100.0%; Score 49; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.00024;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 |||||
 DB 1 YRLAIRLNER 10

RESULT 9

W47268
 ID W47268 standard; peptide; 10 AA.
 AC W47268; (first entry)
 DE Immunomodulatory peptide.
 KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
 KW transplant rejection; treatment; autoimmune disease.
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc_difference 1.10
 FT /note- "at least one of the amino acids is the
 FT D-isomer

PN W09744052-A1.
 PD 27-NOV-1997.
 PF 23-APR-1997; U06705.
 PR 22-MAY-1996; US-651650.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI: 98-018220/02.
 PT Novel immunomodulatory peptide-type compound - useful for inhibiting
 PT transplant rejection
 PS Claim 10; Page 36; 41pp; English.
 CC The present sequence is an immunomodulatory peptide, which
 CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
 CC in a pharmaceutical composition together with a subtherapeutic dose
 CC of an immunosuppressant, to extend the period of acceptance of a
 CC transplant from a major histocompatibility complex (MHC) unmatched
 CC donor, i.e. to inhibit transplant rejection. It can also be used in
 CC the treatment of autoimmune diseases.
 CC Peptides using the D-form amino acids are more effective
 CC immunomodulators than their diastereomers or enantiomers.

SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0013;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 |||||
 DB 1 YRLAIRLNER 10

RESULT 10

W47270
 ID W47270 standard; peptide; 10 AA.
 AC W47270;
 DT 22-MAY-1998 (first entry)
 DE Immunomodulatory peptide.
 KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
 KW transplant rejection; treatment; autoimmune disease.
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc_difference 1.10
 FT /note- "at least one of the amino acids is the
 FT D-isomer

PN W09744052-A1.
 PD 27-NOV-1997.
 PF 23-APR-1997; U06705.
 PR 22-MAY-1996; US-651650.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI: 98-018220/02.
 PT Novel immunomodulatory peptide-type compound - useful for inhibiting
 PT transplant rejection
 PS Claim 10; Page 36; 41pp; English.
 CC The present sequence is an immunomodulatory peptide, which
 CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
 CC in a pharmaceutical composition together with a subtherapeutic dose
 CC of an immunosuppressant, to extend the period of acceptance of a
 CC transplant from a major histocompatibility complex (MHC) unmatched
 CC donor, i.e. to inhibit transplant rejection. It can also be used in
 CC the treatment of autoimmune diseases.
 CC Peptides using the D-form amino acids are more effective
 CC immunomodulators than their diastereomers or enantiomers.

SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0013;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 |||||
 DB 1 YRLAIRLNER 10

RESULT 11

R92910
 ID R92910 standard; peptide; 20 AA.
 AC R92910;
 DT 16-MAY-1996 (first entry)
 DE HLA-B*2702 CTL modulating peptide (B2702.84-75(F)/75-84(T)).
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW Class I MHC; HLA-B*2702.
 OS Synthetic.
 PN W09526979-A1.
 PD 12-OCT-1995.
 PF 05-APR-1995; U04349.
 PR 05-APR-1994; US-222851.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM, Parham P;

DR WPI; 95-358582/46.
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
PS Example 15; Page 36; 80pp; English.
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of
CC class I major histocompatibility complex (MHC) antigens. This sequence
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class
CC I MHC HLA-B2702. These sequences can be used to extend the period of
CC acceptance by a recipient of a transplant from an MHC unmatched donor.
CC The peptides are administered to a patient in conjunction with a
CC subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.
SQ Sequence 20 AA;

Query Match 89.8%; Score 44; DB 1; Length 20;
Best Local Similarity 90.0%; Pred. No. 0.0028;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
| | | | |
DB 1 YRLAIRLNER 10

RESULT 12

DR R92908
ID R92908 standard; peptide; 20 AA.
AC R92908;
AT 16-MAY-1996 (first entry)
DE HLA-B2702 CTL modulating peptide (B2702.84-75(T)/75-84).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW class I MHC; HLA-B2702.
OS Synthetic.
PN W09526979-Al.
PD 12-OCT-1995.
PF 05-APR-1995; U04349.
PR (STRD) UNIV LELAND STANFORD JUNIOR.
PA Clayberger C, Krensky AM, Parham P;
PI WPI: 95-358582/46.
DR Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
PS Example 15; Page 36; 80pp; English.
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of
CC class I major histocompatibility complex (MHC) antigens. This sequence
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class
CC I MHC HLA-B2702. These sequences can be used to extend the period of
CC acceptance by a recipient of a transplant from an MHC unmatched donor.
CC The peptides are administered to a patient in conjunction with a
CC subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.
SQ Sequence 20 AA;

Query Match 89.8%; Score 44; DB 1; Length 20;
Best Local Similarity 90.0%; Pred. No. 0.0028;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
| | | | |
DB 1 YRLAIRLNER 10

RESULT 13

R95430
ID R95430 standard; peptide; 20 AA.
AC R95430;
AT 12-NOV-1996 (first entry)
DE HLA-B2702.84-75(T)/75-84T palindromic.
KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
KW cytolysis; antigen presenting cell.
OS Synthetic.
PN W09513288-Al.
PD 18-MAY-1995.
PF 10-NOV-1994; U12985.
PR 10-NOV-1993; US-150493.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM;
DR WPI: 95-194027/25.
PT Compens. comprising lymphoid surface membrane proteins - which may
PT inhibit cytolytic activity and differentiation of CTLs.
PS Example; Page 12; 29pp; English.
CC R95413, and R95415-R95431 represent palindromes and fragments of
CC human-leucocyte-associated antigens. This sequence represents the
CC HLA-B2702.84-75(T)/75-84T palindromic. These sequences can be used to
CC isolate the protein p74 from a T-cell lysate. p74 is a T-cell surface
CC membrane protein associated with T-cell activation in mammalian T-cells,
CC and is also immunologically cross reactive with the heat shock protein
CC Hsc70. p74 is found in a limited number of cell types, but is
CC particularly expressed on B and T cells. p74 can be isolated by lysis of
CC a suitable cell with an amphoteric detergent, and then passed through an
CC affinity column containing a covalently bound HLA-B2702 palindromic
CC peptide. Compositions comprising the extracellular fragment of p74
CC combined with HLA-B2702.60-84 (see R95416), induces calcium influx, and
CC inhibits cytotoxic T lymphocyte (CTL) differentiation or cytotoxic
CC Candidate compounds can be screened for their effect on the cytolytic
CC activity of T-cells, by combining them with the extracellular portion of
CC p74 and determining the amount of binding between the candidate compound
CC and p74. Modulation of CTL activity can be inhibited in a cellular
CC composition containing T-cells and antigen presenting cells (APCs), by
CC adding to the mix the extracellular portion of p74, in an amount
CC sufficient to compete with p74 for the binding of the p74 ligand.
SQ Sequence 20 AA;

Query Match 89.8%; Score 44; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIRLNE 9
| | | | |
DB 1 YRLAIRLNE 9

RESULT 14

DR W33791
ID W33791 standard; peptide; 20 AA.
AC W33791;
AT 19-JUN-1998 (first entry)
DE Peptide B2702.84-75(T)/75-84 tested for immunomodulating activity.
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
KW rejection.
OS Synthetic.
OS Homo sapiens.
PN W09744351-Al.
PD 27-NOV-1997.
PF 22-MAY-1997; U08689.
PR 24-MAY-1996; US-653294.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Beulow R, Clayberger C, Krensky AM;
DR WPI: 98-086530/08.
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
PT alpha-1 domain, used for preventing rejection of transplants or
PT treating autoimmune diseases

Example 1: Page 19; 41pp; English.

PS Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
CC activity. A peptide-type compound or variant is claimed which has
CC immunomodulating activity, including the N-terminal acylated and/or
CC C-terminal amidated or esterified forms of up to 60 amino acids, where
CC the peptide-type compound comprises the formula: A-B, where A, B =
CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
CC acid. The sequence in the brackets may optionally be absent or truncated
CC at any peptide type bond within the brackets. The compounds comprise
CC amino acid sequences related to a Class I HLA-B alpha domain (positions
CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
CC undesirably attacking cells in a host or in vitro. They can also be
CC used in combination with antigenic peptides or proteins of interest to
CC activate CTLs. They can also inhibit the proliferation of T cells in
CC response to anti-CD3. The peptide can be used for preventing rejection
CC of transplants or for treating autoimmune diseases, e.g. diabetes,
CC rheumatoid arthritis and lupus erythematosus. The products can also be
CC used for detection and diagnosis.
SQ Sequence 20 AA;

Query Match 89.8%; Score 44; DB 1; Length 20;
Best Local Similarity 90.0%; Pred. No. 0.0028;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 |||||
DB 1 YRLATRLNER 10

RESULT 15
W33793

ID W33793 standard; peptide; 20 AA.
AC W33793;
DT 19-JUN-1998 (first entry)
DE Peptide B2702.84-75/75-84r tested for immunomodulating activity.
KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
KW rejection.
OS Synthetic.
OS Homo sapiens.
PN W09744351-A1.
PD 27-NOV-1997.
PF 22-MAY-1997; U08689.
PR 24-MAY-1996; US-653294.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Beulow R, Clayberger C, Krensky AM;
DR WPI; 98-086530/08
PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
PT alpha-1 domain, used for preventing rejection of transplants or
PT treating autoimmune diseases
PS Example 1; Page 19; 41pp; English.
CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
CC activity. A peptide-type compound or variant is claimed which has
CC immunomodulating activity, including the N-terminal acylated and/or
CC C-terminal amidated or esterified forms of up to 60 amino acids, where
CC the peptide-type compound comprises the formula: A-B, where A, B =
CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
CC acid. The sequence in the brackets may optionally be absent or truncated
CC at any peptide type bond within the brackets. The compounds comprise
CC amino acid sequences related to a Class I HLA-B alpha domain (positions
CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
CC undesirably attacking cells in a host or in vitro. They can also be
CC used in combination with antigenic peptides or proteins of interest to
CC activate CTLs. They can also inhibit the proliferation of T cells in
CC response to anti-CD3. The peptide can be used for preventing rejection
CC of transplants or for treating autoimmune diseases, e.g. diabetes,
CC rheumatoid arthritis and lupus erythematosus. The products can also be
CC used for detection and diagnosis.

SQ Sequence 20 AA;

Query Match 89.8%; Score 44; DB 1; Length 20;
Best Local Similarity 90.0%; Pred. No. 0.0028;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 |||||
DB 1 YRLATRLNER 10

Search completed: February 8, 2000, 01:29:37
Job time: 1749 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 11:54:20 ; Search time 117.7 seconds

(without alignments)
4.008 Million cell updates/sec

Title: US-08-653-294-9

Perfect score: 49

Sequence: 1 YRLAIRLNER 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

PIR_62:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 36 | 73.5 | 485 | 1 B45343 | glycoprotein gp13 |
| 2 | 34 | 69.4 | 468 | 1 VGBEEH | glycoprotein gp13 |
| 3 | 34 | 69.4 | 468 | 1 B46114 | glycoprotein gp13 |
| 4 | 34 | 69.4 | 846 | 2 A30889 | integrin beta chain |
| 5 | 33 | 67.3 | 185 | 2 S74416 | hypothetical prote |
| 6 | 33 | 67.3 | 741 | 2 T16992 | ethylene receptor |
| 7 | 33 | 67.3 | 1214 | 2 JC2069 | zinc-finger protei |
| 8 | 32 | 65.3 | 123 | 2 D71922 | probable flagellar |
| 9 | 32 | 65.3 | 123 | 2 H45392 | flagellar switch p |
| 10 | 32 | 65.3 | 465 | 2 JE0369 | histidine acid pho |
| 11 | 32 | 65.3 | 654 | 2 S58820 | DNA-directed RNA p |
| 12 | 32 | 65.3 | 661 | 2 S67177 | hypothetical prote |
| 13 | 32 | 65.3 | 1154 | 2 S43275 | hypothetical prote |
| 14 | 32 | 65.3 | 1154 | 2 S43277 | hypothetical prote |
| 15 | 32 | 65.3 | 1402 | 2 S23557 | hypothetical prote |
| 16 | 32 | 65.3 | 2077 | 1 WBEZ4 | 24kD tegument prot |
| 17 | 32 | 65.3 | 2078 | 2 T09326 | tegument protein - |
| 18 | 31 | 63.3 | 144 | 2 T1047 | ribosomal mobile e |
| 19 | 31 | 63.3 | 182 | 1 TVFER3 | transforming prote |
| 20 | 31 | 63.3 | 321 | 2 C84941 | hypothetical prote |
| 21 | 31 | 63.3 | 377 | 2 D72317 | hypothetical prote |
| 22 | 31 | 63.3 | 395 | 2 T01392 | leucine-rich repea |
| 23 | 31 | 63.3 | 418 | 2 B71041 | probable UDP-N-ace |
| 24 | 31 | 63.3 | 444 | 2 JC6088 | RNA polymerase-ass |
| 25 | 31 | 63.3 | 445 | 2 S44541 | hypothetical prote |
| 26 | 31 | 63.3 | 463 | 2 B72500 | probable seryl-RN |
| 27 | 31 | 63.3 | 538 | 2 D72213 | conserved hypothet |
| 28 | 31 | 63.3 | 584 | 2 S06696 | hypothetical prote |
| 29 | 31 | 63.3 | 624 | 2 T05090 | hypothetical prote |
| 30 | 31 | 63.3 | 738 | 2 A48246 | ethylene-response |

| | | | | | |
|----|----|------|-----|----------|--------------------|
| 31 | 30 | 61.2 | 132 | 1 D69934 | cell wall enzyme h |
| 32 | 30 | 61.2 | 176 | 2 F71321 | conserved hypothet |
| 33 | 30 | 61.2 | 201 | 2 S64994 | probable membrane |
| 34 | 30 | 61.2 | 246 | 2 B25528 | trypsin (EC 3.4.21 |
| 35 | 30 | 61.2 | 252 | 2 H64752 | probable transcrip |
| 36 | 30 | 61.2 | 297 | 2 T12615 | ribosomal protein |
| 37 | 30 | 61.2 | 306 | 2 D70202 | 2-methylthio-N6-is |
| 38 | 30 | 61.2 | 334 | 2 A72217 | Holliday junction |
| 39 | 30 | 61.2 | 336 | 2 H70693 | hypothetical prote |
| 40 | 30 | 61.2 | 345 | 2 H71358 | conserved hypothet |
| 41 | 30 | 61.2 | 377 | 2 A35795 | carbonate dehydrat |
| 42 | 30 | 61.2 | 381 | 2 F71196 | hypothetical prote |
| 43 | 30 | 61.2 | 409 | 2 S74736 | methionine adenosy |
| 44 | 30 | 61.2 | 517 | 2 T11665 | probable cell divi |
| 45 | 30 | 61.2 | 610 | 2 G69130 | conserved hypothet |

ALIGNMENTS

RESULT 1

B45343

glycoprotein gp13 precursor - equine herpesvirus 4

N:Alternate names: glycoprotein gc

C:Species: equine herpesvirus 4

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999

C:Accession: B45343

R:Nicolson, L.; Onions, D.E.

Virol. 179, 378-387, 1990

A:Title: The nucleotide sequence of the equine herpesvirus 4 gc gene homologue.

A:Reference number: A45343; MUID:91021040

A:Accession: B45343

A:Molecule type: DNA

A:Residues: 1-485 <NIC>

A:Cross-references: GB:M58031; NID:g330894; PIDN:AAA46083.1; PID:g330896

C:Genetics:

A:Gene: 16

C:Superfamily: herpesvirus glycoprotein F

C:Keywords: glycoprotein; transmembrane protein

F:1-30/domain: signal sequence #status predicted <SIG>

F:31-485/product: glycoprotein gp13 #status predicted <GGP>

F:60,61,66,67,72,108,116,147,220,225,286/Binding site: carbohydrate (Asn) (covalent)

Query Match 73.5%; Score 36; DB 1; Length 485;

Best Local Similarity 70.0%; Pred. No. 6.1;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10

DB 120 YRLTEHLNQR 129

RESULT 2

VGBEEH

glycoprotein gp13 precursor - equine herpesvirus 1

N:Alternate names: glycoprotein gc

C:Species: equine herpesvirus 1

A:Note: host Equus caballus (domestic horse)

C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 24-Sep-1999

C:Accession: A28149; A32980; H36796

R:Allen, G.P.; Coogle, L.D.

J. Virol. 62, 2850-2858, 1988

A:Title: Characterization of an equine herpesvirus type 1 gene encoding a glycoprotein

A:Reference number: A28149; MUID:88275055

A:Accession: A28149

A:Molecule type: DNA

A:Residues: 1-468 <ALL>

A:Cross-references: GB:M86664; NID:g330791; PIDN:AAB02451.1; PID:g330808

A:Experimental source: strain Kentucky T431

A:Note: the authors translated the codon ACA for residue 43 as Pro

R:Guo, P.; Goebel, S.; Davis, S.; Perkus, M.E.; Languet, B.; Desmettre, P.; Allen, G.

J. Virol. 63, 4189-4198, 1989

A:Title: Expression in recombinant vaccinia virus of the equine herpesvirus 1 gene encoded
 A:Reference number: A32980; MUID:93382761
 A:Accession: A32980
 A:Molecule type: DNA
 A:Residues: 1-468 <G00>
 A:Cross-references: GB:M86664; NID:g330791; PIDN:AB02451.1; PID:g330808
 R:Teiford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
 submitted to GenBank, March 1992
 A:Description: The DNA sequence of equine herpesvirus-1.
 A:Reference number: A36805
 A:Accession: H36796
 A:Molecule type: DNA
 A:Residues: 1-468 <TEL>
 A:Cross-references: GB:M86664; NID:g330791; PIDN:AB02451.1; PID:g330808
 A:Experimental source: strain Ab4P
 R:Teiford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
 Virology 189, 304-316, 1992
 A:Title: The DNA sequence of equine herpesvirus-1.
 A:Reference number: A41831; MUID:92295566
 A:Contents: annotation; possible protein-coding frames
 A:Note: neither amino acid nor nucleotide sequence is given
 C:Genetics:
 A:Gene: 16
 C:Superfamily: herpesvirus glycoprotein F
 C:Keywords: glycoprotein; transmembrane protein
 F:1-30/Domain: signal sequence #status predicted <SIG>
 F:31-468/Product: glycoprotein gp13 #status predicted <MAT>
 F:46,57,62,92,100,131,203,208,269/Binding site: carbohydrate (Asn) (covalent) #status pr

Query Match 69.4%; Score 34; DB 1; Length 468;
 Best Local Similarity 70.0%; Pred. No. 16;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 ||| | |||
 Db 104 YRLKIYLNQR 113

RESULT 3
 B46114
 glycoprotein gp13 precursor - equine herpesvirus 1 (strain Kentucky A)
 N:Alternate names: glycoprotein C
 C:Species: equine herpesvirus 1
 A:Note: host Equus caballus (domestic horse)
 C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
 C:Accession: B46114
 R:Matsumura, T.; Smith, R.H.; O'Callaghan, D.J.
 Virology 193, 910-923, 1993
 A:Title: DNA sequence and transcriptional analyses of the region of the equine herpesvir

A:Reference number: A46114; MUID:93212524
 A:Accession: B46114
 A:Molecule type: DNA
 A:Residues: 1-468 <MAT>
 A:Cross-references: GB:S57839; NID:g298846; PIDN:AB025944.1; PID:g298848
 C:Superfamily: herpesvirus glycoprotein F
 C:Keywords: glycoprotein; transmembrane protein
 F:1-30/Domain: signal sequence #status predicted <SIG>
 F:31-468/Product: glycoprotein gp13 #status predicted <GPT>
 F:432-451/Domain: transmembrane #status predicted <TMN>
 F:46,57,62,92,100,131,203,208,269/Binding site: carbohydrate (Asn) (covalent) #status pr

Query Match 59.4%; Score 34; DB 1; Length 468;
 Best Local Similarity 70.0%; Pred. No. 16;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 ||| | |||
 Db 104 YRLKIYLNQR 113

RESULT 4

A30889
 integrin beta chain precursor - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 20-Aug-1999
 C:Accession: A30889
 R:MacKrell, A.J.; Blumberg, B.; Haynes, S.R.; Fessler, J.H.
 Proc. Natl. Acad. Sci. U.S.A. 85, 2633-2637, 1988
 A:Title: The lethal myospheroid gene of Drosophila encodes a membrane protein homolog
 A:Reference number: A30889; MUID:88190122
 A:Accession: A30889
 A:Molecule type: mRNA
 A:Residues: 1-846 <NAC>
 A:Cross-references: GB:J03251; NID:g157954; PIDN:AAA28714.1; PID:g157955
 C:Genetics:
 A:Gene: FlyBase:mys
 A:Cross-references: FlyBase:FBgn0004657
 C:Superfamily: integrin beta chain; laminin-type EGF-like homology
 C:Keywords: cell adhesion; cytoskeleton; transmembrane protein

Query Match 69.4%; Score 34; DB 2; Length 846;
 Best Local Similarity 66.7%; Pred. No. 29;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLAIRLNER 10
 ||| | |||
 Db 162 RLALRVNEK 170

RESULT 5
 S74416
 hypothetical protein sll0687 - Synecocystis sp. (strain PCC 6803)
 C:Species: Synecocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 21-Aug-1998
 C:Accession: S74416
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocysts

A:Reference number: S74322; MUID:97061201
 A:Accession: S74416
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-185 <KAN>
 A:Cross-references: EMBL:D64001; GB:AB001339; NID:g1001102; PID:d1010985; PID:g100119
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 67.3%; Score 33; DB 2; Length 185;
 Best Local Similarity 60.0%; Pred. No. 9.5;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 ||| | |||
 Db 35 YRLALRILOL 44

RESULT 6
 T16992
 ethylene receptor homolog - apple tree
 C:Species: Malus domestica (apple tree)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T16992
 R:Chen, H.H.; Charing, Y.Y.; Yang, S.F.; Shaw, J.F.
 Plant Physiol. 117, 1126, 1998
 A:Title: Isolation and characterization of a broccoli cDNA (Accession No. AF047477) e
 A:Reference number: Z18085
 A:Accession: T16992
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-741 <LEE>

A:Cross-references: EMBL:AF032448; NID:g3411050; PID:g3411051
A:Experimental source: cultivar Granny Smith; ripening fruit
C:Genetics:
A:Note: ETR1

Query Match 67.3%; Score 33; DB 2; Length 741;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
| | | | |
Db 672 YELAVRIHEK 681

RESULT 7
JC2069
Zinc-finger protein, BR140 - human
N:Alternate names: bromodomain protein
C:Species: Homo sapiens (man)
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 12-Sep-1997
C:Accession: JC2069
R:Thompson, K.A.; Wang, B.; Argraves, W.S.; Giannocotti, F.G.; Schranck, D.P.; Ruoslahti, Biochem. Biophys. Res. Commun. 198, 1143-1152, 1994
A:Title: BR140, a novel zinc-finger protein with homology to the TAF250 subunit of TFIID
A:Reference number: JC2069; MUID:94161726
A:Accession: JC2069
A:Molecule type: mRNA
A:Residues: 1-1214 <THO>
A:Cross-references: GB:M91585
C:Comment: This is a nuclear protein with broad tissue distribution, but is especially a C:Superfamily: unassigned bromodomain proteins; bromodomain homology
C:Keywords: DNA binding; phosphoprotein; transcription regulation; zinc finger
F:653-708/Domain: bromodomain homology <BRO>
F:23-28,41,44/Binding site: zinc (Cys, Cys, His, His) #status predicted
F:120,205,462/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status pre
F:276,279,293,296/Binding site: zinc (Cys, Cys, Cys, Cys) #status predicted
F:301,304,317,320/Binding site: zinc (His, Cys, Cys, Cys) #status predicted
F:330,333,350,353/Binding site: zinc (Cys, Cys, His, Cys) #status predicted
F:386,389,401,405/Binding site: zinc (Cys, Cys, His, Cys) #status predicted
F:410,413,444,447/Binding site: zinc (His, Cys, Cys, His) #status predicted

Query Match 67.3%; Score 33; DB 2; Length 1214;
Best Local Similarity 60.0%; Pred. No. 70;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
| | | | |
Db 715 YRAAVRLREQ 724

RESULT 8
D71922
Probable flagellar motor switch protein - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 26-Aug-1999
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557
A:Accession: D71922
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-123 <ARN>
A:Cross-references: GB:AE001485; GB:AE001439; NID:g4155057; PIDN:AAD06097.1; PID:g415506
A:Experimental source: strain J99
C:Genetics:
A:Gene: fln
C:Superfamily: flagellar motor switch protein

Query Match 65.3%; Score 32; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAIRLNE 9
| | | | |
Db 102 LAIRLNE 108

RESULT 9
H64592
Flagellar switch protein - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 26-Aug-1999
C:Accession: H64592
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R. Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKee son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467
A:Accession: H64592
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-123 <TOM>
A:Cross-references: GB:AE000571; GB:AE000511; NID:g2313686; PIDN:AAD07636.1; PID:g231 C:Superfamily: flagellar motor switch protein

Query Match 65.3%; Score 32; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAIRLNE 9
| | | | |
Db 102 LAIRLNE 108

RESULT 10
JE0369
Histidine acid phosphatase (EC 3.1.3.-) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 20-Sep-1999
C:Accession: JE0369
R:Mullaney, E.J.; Ullah, A.H.J.
A:Title: Identification of a histidine acid phosphatase (phyA)-like gene in Arabidopsis Biochem. Biophys. Res. Commun. 251, 252-255, 1998
A:Title: Identification of a histidine acid phosphatase
A:Reference number: JE0369
A:Accession: JE0369
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-465 <MWL>
C:Superfamily: yeast acid phosphatase
C:Keywords: phosphoric monoester hydrolase

Query Match 65.3%; Score 32; DB 2; Length 465;
Best Local Similarity 60.0%; Pred. No. 41;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
| | | | |
Db 128 YOLGIRVRER 137

RESULT 11
S58820
DNA-directed RNA polymerase (EC 2.7.7.6) III chain RPC92 - yeast (Saccharomyces cerev N:Alternate names: DNA-directed RNA polymerase C chain RPC82; DNA-directed RNA poly C:Species: Saccharomyces cerevisiae

C:Date: 28-Nov-1995 #sequence_revision 19-Jan-1996 #text_change 23-Apr-1999
C:Accession: S58820; S31298

R:Miller, N.
submitted to the EMBL Data Library, April 1995
A:Description: The sequence of S. cerevisiae cosmid 9677.
A:Reference number: S58816
A:Accession: S58820
A:Molecule type: DNA
A:Residues: 1-654 <MIL>
A:Cross-references: EMBL:U25841; NID:g786295; PID:g786307; MIPS:YPR190c
R:Chlannilkulchai, N.; Stalder, R.; Riva, M.; Carles, C.; Werner, M.; Sentenac, A.
Mol. Cell. Biol. 12, 4433-4440, 1992
A:Title: RCB82 encodes the highly conserved, third-largest subunit of RNA polymerase C
A:Reference number: S31298; MUID:93024385

A:Accession: S31298
A:Molecule type: DNA
A:Residues: 1-636, 'L', 638-654 <CHI>
A:Cross-references: EMBL:X63500; NID:g4383; PID:g4384
C:Genetics:
A:Gene: SGD:RPC82
A:Cross-references: SGD:S0006394; MIPS:YPR190c
A:Map position: 16R
C:Keywords: leucine zipper; nucleotidyltransferase; transcription

Query Match 65.3%; Score 32; DB 2; Length 654;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
|||:|:|:|
Db 312 YKIALRLTEQ 321

RESULT 12

S67177
hypothetical protein YOR275c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O5450
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 26-Feb-1998
C:Accession: S67177; S72046
R:Cheret, G.; Sor, F.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67169
A:Accession: S67177

A:Molecule type: DNA
A:Residues: 1-661 <CHE>
A:Cross-references: EMBL:Z75183; NID:g1420615; PID:e252121; PID:g1420616; MIPS:YOR275c
A:Experimental source: strain S288C
R:Cheret, G.; Bernardi, A.; Sor, F.
Yeast 12, 1059-1064, 1996
A:Title: DNA sequence analysis of the VPB1-SNF2 region on chromosome XV of Saccharomyces
A:Reference number: S72039; MUID:97051594

A:Accession: S72046
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-661 <CHW>
A:Cross-references: EMBL:X89633; NID:g1279694; PID:e189401; PID:g1279701
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995
C:Genetics:
A:Map position: 15R
A:Note: YOR275c

Query Match 65.3%; Score 32; DB 2; Length 661;
Best Local Similarity 50.0%; Pred. No. 60;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
|||:|:|:|
Db 241 YRIALSFNEK 250

RESULT 13

S43275
hypothetical protein 2 - Neurospora crassa retrotransposon Tad1-1
C:Species: Neurospora crassa
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Sep-1997
C:Accession: S43275
R:Cambareri, E.B.; Helber, J.; Kinsey, J.A.

Mol. Gen. Genet. 242, 658-665, 1994
A:Title: Tad1-1, an active LINE-like element of Neurospora crassa.
A:Reference number: S43274; MUID:94203179

A:Accession: S43275
A:Molecule type: DNA
A:Residues: 1-1154 <CAN>
A:Cross-references: EMBL:L25662; NID:g409759; PID:g409761
C:Genetics:
A:Mobile element: retrotransposon Tad1-1

Query Match 65.3%; Score 32; DB 2; Length 1154;
Best Local Similarity 66.7%; Pred. No. 11e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNE 9
||||:|
Db 1136 YRLAVELEE 1144

RESULT 14

S43277
hypothetical protein 2 - Neurospora crassa retrotransposon Tad3-2
C:Species: Neurospora crassa
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Sep-1997
C:Accession: S43277
R:Cambareri, E.B.; Helber, J.; Kinsey, J.A.

Mol. Gen. Genet. 242, 658-665, 1994
A:Title: Tad1-1, an active LINE-like element of Neurospora crassa.
A:Reference number: S43274; MUID:94203179

A:Accession: S43277
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1154 <CAN>
A:Cross-references: EMBL:L25663; NID:g409762; PID:g409764
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993

Query Match 65.3%; Score 32; DB 2; Length 1154;
Best Local Similarity 66.7%; Pred. No. 11e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNE 9
||||:|
Db 1136 YRLAVELEE 1144

RESULT 15

S62557
hypothetical protein SPAC24B11.12c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 16-May-1996 #sequence_revision 13-Mar-1997 #text_change 31-Oct-1997
C:Accession: S62557
R:Odell, C.; Churcher, C.M.

submitted to the EMBL Data Library, November 1995
A:Reference number: S62546
A:Accession: S62557
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1402 <CODE>
A:Cross-references: EMBL:S67757; NID:g1061288; PID:g1061300
C:Genetics:
A:Map position: 1L
A:Introns: 124/3

Query Match 65.3%; Score 32; DB 2; Length 1402;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLA1RLN 8
DB 25 YRLA1RLN 32

Search completed: February 7, 2000, 11:54:22
Job time: 24332 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 00:59:49 ; Search time 63.71 Seconds
(without alignments)
4.688 Million cell updates/sec

Title: US-08-653-294-9
Perfect score: 49
Sequence: 1 YRLAIRLNER 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 82229 seqs, 29864866 residues

Total number of hits satisfying chosen parameters: 82229

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SwissProt_38:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 36 | 73.5 | 485 | 1 | VLGLC_HSV4 |
| 2 | 34 | 69.4 | 468 | 1 | VLGLC_HSV4 |
| 3 | 34 | 69.4 | 846 | 1 | ITBX_DROME |
| 4 | 33 | 67.3 | 1214 | 1 | BR14_HUMAN |
| 5 | 32 | 65.3 | 654 | 1 | RPC3_YEAST |
| 6 | 32 | 65.3 | 1402 | 1 | ATCX_SCHPO |
| 7 | 32 | 65.3 | 2077 | 1 | TEGU_HSV6G |
| 8 | 32 | 65.3 | 2077 | 1 | TEGU_HSV6G |
| 9 | 31 | 63.3 | 321 | 1 | YKAX_ECOLI |
| 10 | 31 | 63.3 | 445 | 1 | PAF1_YEAST |
| 11 | 31 | 63.3 | 584 | 1 | 65KD_ZYMO |
| 12 | 31 | 63.3 | 738 | 1 | ETRL_ARATH |
| 13 | 30 | 61.2 | 132 | 1 | YPOQ_BACSU |
| 14 | 30 | 61.2 | 246 | 1 | TRIP_MOUSE |
| 15 | 30 | 61.2 | 252 | 1 | YAGI_ECOLI |
| 16 | 30 | 61.2 | 294 | 1 | RL5A_SCHPO |
| 17 | 30 | 61.2 | 294 | 1 | RL5B_SCHPO |
| 18 | 30 | 61.2 | 297 | 1 | RL5_HELAN |
| 19 | 30 | 61.2 | 299 | 1 | RL5_BOMMO |
| 20 | 30 | 61.2 | 306 | 1 | MAAA_BORBU |
| 21 | 30 | 61.2 | 334 | 1 | RUVB_THEMA |
| 22 | 30 | 61.2 | 340 | 1 | AC11_HUMAN |
| 23 | 30 | 61.2 | 377 | 1 | CAH1_CHLRE |
| 24 | 30 | 61.2 | 409 | 1 | METK_SYNY3 |
| 25 | 30 | 61.2 | 894 | 1 | RN6_YEAST |
| 26 | 29 | 59.2 | 211 | 1 | BAK2_HUMAN |
| 27 | 29 | 59.2 | 211 | 1 | IPYR_SOLTU |
| 28 | 29 | 59.2 | 215 | 1 | PCP_BACAM |
| 29 | 29 | 59.2 | 232 | 1 | CYSH_SYNP7 |
| 30 | 29 | 59.2 | 267 | 1 | DPM1_YEAST |
| 31 | 29 | 59.2 | 301 | 1 | RL5_NEUCR |
| 32 | 29 | 59.2 | 327 | 1 | RL5_ANOGA |
| 33 | 29 | 59.2 | 387 | 1 | YGBA_BACSU |
| 34 | 29 | 59.2 | 517 | 1 | NCAP_SENDE |

35 29 59.2 517 1 NCAP_SENDZ P04858 sendai viru
36 29 59.2 524 1 NCAP_SEND5 P27563 sendai viru
37 29 59.2 524 1 NCAP_SENDF Q07097 sendai viru
38 29 59.2 525 1 LAG3_HUMAN P18627 homo sapien
39 29 59.2 525 1 Y107_METJA Q57571 methanococc
40 29 59.2 537 1 MYPH_CHICK Q05623 gallus gall
41 29 59.2 569 1 MYPE_ECOLI P16431 escherichia
42 29 59.2 645 1 SKB1_SCHPO P78963 schizosacch
43 29 59.2 665 1 NUC2_SCHPO P10505 schizosacch
44 29 59.2 685 1 CRPI_PERAM Q25641 periplaneta
45 29 59.2 759 1 MCM5_CAEEL Q21902 caenorhabdi

ALIGNMENTS

RESULT 1
VLGLC_HSV4
ID VLGLC_HSV4 STANDARD; PRT; 485 AA.
AC P22596;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCOPROTEIN C PRECURSOR (GLYCOPROTEIN 13).
GN GC OR GP13.
OS Equine herpesvirus type 4 (strain 1942) (EHV-4) (Equine herpesvirus
OS type 1 subtype 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91021040.
RA NICOLSON L., ONIONS D.E.;
RT "The nucleotide sequence of the equine herpesvirus 4 GC gene
RT homologue.";
RL Virology 179:378-387(1990).
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN C.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN GENE SUPERFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL; M58031; AAA46083.1; -;
CC EMBL; A21044; CAA01528.1; -;
CC PIR; B45343; B45343.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 32
FT CHAIN 33 485 GLYCOPROTEIN C.
FT DOMAIN 33 444 EXTRACELLULAR.
FT TRANSMEM 445 468
FT CARBOHYD 60 60 POTENTIAL.
FT CARBOHYD 61 61 POTENTIAL.
FT CARBOHYD 66 66 POTENTIAL.
FT CARBOHYD 67 67 POTENTIAL.
FT CARBOHYD 72 72 POTENTIAL.
FT CARBOHYD 108 108 POTENTIAL.
FT CARBOHYD 116 116 POTENTIAL.
FT CARBOHYD 147 147 POTENTIAL.
FT CARBOHYD 220 220 POTENTIAL.
FT CARBOHYD 225 225 POTENTIAL.
FT CARBOHYD 286 286 POTENTIAL.
SQ SEQUENCE 485 AA; 52509 MW; 63F72464 CRC32;

Query Match 73.5%; Score 36; DB 1; Length 485;
Best Local Similarity 70.0%; Pred. No. 3.2;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 DB 120 YRLHLNR 129

RESULT 2
 VGLC_HSVB STANDARD; PRT; 468 AA.

AC P12889; P36321;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GLYCOPROTEIN C PRECURSOR (GLYCOPROTEIN 13).
 GN GC OR gp13 OR 16.
 OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1), and
 OS Equine herpesvirus type 1 (strain Kentucky D) (EHV-1).
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AB4P;
 RX MEDLINE; 92295566.
 RA TELFORD E.A.R., WATSON M.S., MCBRIDE K., DAVISON A.J.;
 RT "The DNA sequence of equine herpesvirus-1.";
 RL Virology 189:304-316(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KENTUCKY D;
 RX MEDLINE; 88275055.
 RA ALLEN G.P., COOGLE L.D.;
 RT "Characterization of an equine herpesvirus type 1 gene encoding a
 RT glycoprotein (gp13) with homology to herpes simplex virus
 RT glycoprotein C.";
 RL J. Virol. 62:2850-2858(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KENTUCKY D;
 RX MEDLINE; 89382761.
 RA GUO P., GOBEL S., DAVIS S., PERKUS M.E., LANGUET B., DESMETTRE P.,
 RA ALLEN G., PAOLETTI E.;
 RT "Expression in recombinant vaccinia virus of the equine herpesvirus 1
 RT gene encoding glycoprotein gp13 and protection of immunized
 RT animals.";
 RL J. Virol. 63:4189-4198(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KENTUCKY D;
 RX MEDLINE; 93212524.
 RA MATSUMURA T., SMITH R.H., O'CALLAGHAN D.J.;
 RT "DNA sequence and transcriptional analyses of the region of the
 RT equine herpesvirus type 1 Kentucky A strain genome encoding
 RT glycoprotein C.";
 RL Virology 193:910-923(1993).
 CC -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN C.
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN GENE SUPERFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L07272; AAA46078.1;
 CC EMBL; M86664; AAB02451.1;
 CC EMBL; M19966; AAA46077.1;
 CC EMBL; M29234; AAA46085.1;
 CC EMBL; S57839; AAB25944.1;
 CC PIR; H36796; VGBEAL.
 CC PIR; A28149; VGBEEH.
 CC PIR; B46114; B46114.
 CC Glycoprotein; Transmembrane; Signal.

QY 1 YRLAIRLNER 10
 DB 104 YRLHLNR 113

RESULT 3
 ITBX_DROME STANDARD; PRT; 846 AA.

AC P11584;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE INTEGRIN BETA-SUBUNIT MYOSPHEROID PRECURSOR.
 GN L(1)MYS.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 88190122.
 RA MACKRELL A.J., BLUMBERG B., HAYNES S.R., FESSLER J.H.;
 RT "The lethal myospheroid gene of Drosophila encodes a membrane protein
 RT homologous to vertebrate integrin beta subunits.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:2633-2637(1988).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE; 94163982.
 RA GRINBLAT Y., ZUSMAN S., YEE G., HYNES R.O., KAFATOS F.C.;
 RT "Functions of the cytoplasmic domain of the beta PS integrin subunit
 RT during Drosophila development.";
 RL Development 120:91-102(1994).
 CC -!- FUNCTION: NOT KNOWN. PROBABLY PLAYS A ROLE IN CELL ADHESION.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J03251; AAA28714.1;
 CC PIR; A30889; A30889.
 CC FLYBASE; FBgn0004657; mys.
 CC PROSITE; PS00243; INTEGRIN_BETA; 3.
 CC PROSITE; PS00022; EGF_1; UNKNOWN_2.
 CC PROSITE; PS01186; EGF_2; UNKNOWN_1.
 CC PFAM; PF00362; Integrin_B; 1.

Query Match 69.4%; Score 34; DB 1; Length 468;
 Best Local Similarity 70.0%; Pred. No. 8.1;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

KW Integrin; Cell adhesion; Transmembrane; Glycoprotein; Repeat;
 KW Extracellular matrix; Cytoskeleton; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 846 INTEGRIN BETA-SUBUNIT MYOSPIREROID.
 FT DOMAIN 24 776 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 777 799 POTENTIAL.
 FT DOMAIN 800 846 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 115 143 SER-RICH.
 FT DOMAIN 507 687 CYSTEINE-RICH REPEATS.
 FT REPEAT 507 560 I.
 FT REPEAT 561 605 II.
 FT REPEAT 606 646 III.
 FT REPEAT 647 687 IV.
 FT CARBOHYD 72 72 POTENTIAL.
 FT CARBOHYD 266 266 POTENTIAL.
 FT CARBOHYD 277 277 POTENTIAL.
 FT CARBOHYD 403 403 POTENTIAL.
 FT CARBOHYD 428 428 POTENTIAL.
 FT CARBOHYD 557 557 POTENTIAL.
 FT CARBOHYD 603 603 POTENTIAL.
 FT CARBOHYD 644 644 POTENTIAL.
 FT CARBOHYD 718 718 POTENTIAL.
 SQ SEQUENCE 846 AA; 92687 MW; 9906C2F9 CRC32;

Query Match 69.4%; Score 34; DB 1; Length 846;
 Best Local Similarity 66.7%; Pred. No. 16;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 RLAIRLNER 10
 |||:|:|:
 DB 162 RLALRVNEK 170

RESULT 4

BR14_HUMAN
 ID BR14_HUMAN STANDARD; PRT; 1214 AA.
 AC P55201.
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE PEREGRIN (BR140 PROTEIN).
 GN BR140.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 94161726.
 RA THOMPSON K.A., WANG B., ARGRAVES W.S., GIANCOTTI F.G., SCHRANCK D.P.,
 RA RUOSLAHTI E.;
 RT "BR140, a novel zinc-finger protein with homology to the TAF250
 subunit of TFIID.";
 RL Biochem. Biophys. Res. Commun. 198;1143-1152(1994).
 CC -1- FUNCTION: UNKNOWN. POSSIBLE TRANSCRIPTION ACTIVATOR.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
 CC -1- TISSUE SPECIFICITY: HIGH LEVELS IN TESTIS.
 CC -1- SIMILARITY: CONTAINS 1 BROMODOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M91585; AB02119.1;
 DR MIM; 602410;
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 1.
 DR PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
 DR PROSITE; PS00144; BROMODOMAIN_2; 1.
 DR PFAM; PF00439; bromodomain; 1.

DR PFAM; PF00628; PHD; 1.
 KW Transcription regulation; DNA-binding; Activator; Nuclear protein;
 KW Zinc-finger.
 FT ZN_FING 21 47 C2H2-TYPE.
 FT ZN_FING 276 296 C4-TYPE.
 FT ZN_FING 317 333 C4-TYPE.
 FT ZN_FING 386 400 C4-TYPE.
 FT DOMAIN 645 715 BROMODOMAIN.
 SQ SEQUENCE 1214 AA; 137542 MW; B3E44584 CRC32;

Query Match 67.3%; Score 33; DB 1; Length 1214;
 Best Local Similarity 60.08; Pred. No. 38;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 YRLAIRLNER 10
 |||:|:|:
 DB 715 YRAAVRLREQ 724

RESULT 5

RPC3_YEAST
 ID RPC3_YEAST STANDARD; PRT; 654 AA.
 AC P32349; O06591.
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DNA-DIRECTED RNA POLYMERASE III 74 KD POLYPEPTIDE (EC 2.7.7.6) (C74).
 GN RPC3 OR RPC82 OR YPR190C OR P9677.11.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE: 93024385.
 RA CHIANNILKULCHAI N., STALDER R., RIVA M., CARLES C., WERNER M.,
 RA SENTENAC A.;
 RT "RPC82 encodes the highly conserved, third-largest subunit of RNA
 polymerase C (III) from Saccharomyces cerevisiae.";
 RL Mol. Cell. Biol. 12:4433-4440(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,
 RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J., KUCABA T.,
 RA HALLSWORTH K., HAWKINS J., HILLIER L., JIER M., JOHNSON D.,
 RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARDIS E., MENEZES S.,
 RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIFKEN L., RILES L.,
 RA TAICH A., TREVASKIS E., VIGNATI D., WILCOX L., WOHLDMAN P., VAUDIN M.,
 RA WILSON R., WATERSTON R.;
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE +
 CC RNA(N).
 CC -1- SUBUNIT: RNA POLYMERASE III CONSISTS OF ABOUT 15 DIFFERENT
 CC SUBUNITS. THIS SUBUNIT IS THE THIRD LARGEST COMPONENT OF RNA
 CC POLYMERASE III.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE
 CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
 CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
 CC III FOR 5S AND TRNA GENES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

```
-----
CC EMBL; X63500; CAA45072.1; -.
DR EMBL; U25841; AAB64619.1; -.
DR PIR; S31298; S31298.
DR SGD; L0001693; RPC82.
KW Transferase: DNA-directed RNA polymerase; Transcription; Zinc;
Nuclear protein.
FT DOMAIN 581 602 LEUCINE-ZIPPER.
FT CONFLICT 637 637 V -> L (IN REF. 1).
SQ SEQUENCE 654 AA; 74016 MW; 9E17F4F8 CRC32;
-----
Query Match 65.3%; Score 32; DB 1; Length 654;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 YRLAIRLNER 10
DB 312 YKIALRLTEQ 321
-----
RESULT 6
ATCX_SCHPO STANDARD; PRT; 1402 AA.
ID ATCX_SCHPO STANDARD; PRT; 1402 AA.
AC Q09891;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROBABLE CALCIUM-TRANSPORTING ATPASE (EC 3.6.1.38).
GN SPAC24B11.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
RN RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RA SRAIN-972;
RA ODELL C., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RA Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RL FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE
CC HYDROLYSIS OF ATP COUPLED WITH THE TRANSPORT OF CALCIUM
CC (BY SIMILARITY).
CC (-) CATALYTIC ACTIVITY: ATP + H2O = ADP + ORTHOPHOSPHATE.
CC (-) SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN
CC (-) SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC (E1-E2 ATPASES).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL; 267757; CAA91777.1; -.
DR PROSITE; PS00154; ATPASE_E1_E2; 1.
DR PFAM; PF00122; E1-E2_ATPase; 2.
KW Hydrolyase; Calcium transport; Transmembrane; Phosphorylation;
KW Magnesium; ATP-binding.
FT TRANSMEM 109 129 POTENTIAL.
FT TRANSMEM 135 155 POTENTIAL.
FT TRANSMEM 457 477 POTENTIAL.
FT TRANSMEM 501 521 POTENTIAL.
FT TRANSMEM 1066 1086 POTENTIAL.
FT TRANSMEM 1101 1121 POTENTIAL.
FT TRANSMEM 1151 1171 POTENTIAL.
FT TRANSMEM 1193 1213 POTENTIAL.
FT TRANSMEM 1218 1238 POTENTIAL.
FT TRANSMEM 1260 1280 POTENTIAL.
FT MOD_RES 569 569 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 1402 AA; 159355 MW; DEEDAD76 CRC32;
-----
Query Match 65.3%; Score 32; DB 1; Length 1402;
Best Local Similarity 87.5%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 YRLAIRLN 8
DB 25 YRLADRLN 32
-----
RESULT 7
TEGU_HSV6G STANDARD; PRT; 2077 AA.
ID TEGU_HSV6G STANDARD; PRT; 2077 AA.
AC P30002;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE LARGE TEGUMENT PROTEIN.
GN U31.
OS Herpes simplex virus (type 6 / strain GS).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Roseolovirus.
RN RN SEQUENCE FROM N.A.
RP MEDLINE; 91374623.
RA JOSEPHS S.F., ABLASHI D.V., SALAHUDDIN S.Z., JAGODZINSKI L.L.,
RA WONG-STAAAL F., GALLO R.C.;
RA "Identification of the human herpesvirus 6 glycoprotein H and
RT putative large tegument protein genes.";
RL J. Virol. 65:5597-5604(1991).
CC (-) FUNCTION: TEGUMENT PROTEIN.
CC (-) SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,
CC HSV-1 24, EBV BPLF1, HVS-1 64, VZV 22, AND HCMV UL48.
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL; S57540; AAB19786.1; -.
DR PIR; H40511; WZBE24.
SQ SEQUENCE 2077 AA; 239909 MW; 82ACA5DE CRC32;
-----
Query Match 65.3%; Score 32; DB 1; Length 2077;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 YRLAIRLNER 10
DB 495 YKVAVLLNEK 504
-----
RESULT 8
TEGU_HSV6U STANDARD; PRT; 2077 AA.
ID TEGU_HSV6U STANDARD; PRT; 2077 AA.
AC P52340;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE LARGE TEGUMENT PROTEIN.
GN U31 OR HHRF1.
OS Herpes simplex virus (type 6 / strain Uganda-1102).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Roseolovirus.
RN RN SEQUENCE FROM N.A.
RP MEDLINE; 95266321.
RA GOMPELS U.A., NICHOLAS J., LAWRENCE G., JONES M., THOMSON B.J.,
RA MARTIN M.E., EFSTATHIOU S., CRAXTON M., MACAULAY H.A.;
```


RT "The DNA sequence of human herpesvirus-6: structure, coding content,
 RL and genome evolution.";
 CC Virology 209:29-51(1995).
 CC -!- FUNCTION: SEGMENT PROTEIN.
 CC EHV-1 24, EBV BPLF1, HVS-1 64, VZV 22, AND HCMV UL48.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X83413; CA558411.1; -;
 CC SEQUENCE 2077 AA; 239946 MW; 36FCF7B1 CRC32;
 CC -----
 CC Query Match 65.3%; Score 32; DB 1; Length 2077;
 CC Best Local Similarity 50.0%; Pred. No. 1.1e+02;
 CC Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC QY 1 YRLAIRLNER 10
 CC |:::|::|
 CC Db 495 YKVAVLNLEK 504
 CC -----
 CC RESULT 9
 CC ID YEAX_ECOLI STANDARD; PRT; 321 AA.
 CC AC P76254; O07972; O07970;
 CC DT 15-JUL-1998 (Rel. 36, Created)
 CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
 CC DE PUTATIVE DIOXYGENASE BETA SUBUNIT YEAX (EC 1.-.-.-).
 CC GN YEAX.
 CC OS Escherichia coli.
 CC OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC EC Escherichia.
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-K12 / MG1655;
 CC RX MEDLINE: 97426617
 CC RA BLATNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
 CC RILEY M., COLLADO-VIDES J., GLASNER F.D., ROSE C.K., MAYHEW G.F.,
 CC GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
 CC MAU B., SHAO Y.;
 CC "The complete genome sequence of Escherichia coli K-12.";
 CC Science 277:1453-1474(1997).
 CC [2]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-K12;
 CC RX MEDLINE: 97251358.
 CC RA ITOH T., ALBA H., BABA T., FUJITA K., HAYASHI K., INADA T., ISONO K.,
 CC KASAI H., KIMURA S., KITAKAWA M., KITAGAWA M., MAKINO K., MIKI T.,
 CC MIZOBUCHI K., MORI H., MORI T., MOTOMURA K., NAKADE S., NAKAMURA Y.,
 CC NASHIMOTO H., NISHIO H., OSHIMA T., SAITO N., SAMPEI G., SEKI Y.,
 CC SIVASUNDARAM S., TAGAMI H., TAKEDA J., TAKEMOTO K., WADA C.,
 CC YAMAMOTO Y., HORIUCHI T.;
 CC "A 460-kb DNA sequence of the Escherichia coli K-12 genome
 CC corresponding to the 40.1-50.0 min region on the linkage map.";
 CC DNA Res. 3:379-392(1996).
 CC -!- COFACTOR: FMN (BY SIMILARITY).
 CC -!- SUBUNIT: PROBABLE HETERODIMER OF YEAW AND YEAX.
 CC -!- SIMILARITY: IN THE C-TERMINAL, BELONGS TO THE 2FE2S PLANT-TYPE
 CC FERREDOXIN FAMILY.
 CC -!- SIMILARITY: BELONGS TO THE PDR/VANE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE000274; AAC74873.1; -;
 CC EMBL: D90824; CAB21531.1; -;
 CC EMBL: D90823; CAB21524.1; -;
 CC HSSP: P33164; 2PIA.
 CC DR EMOGENE: EG13510; YEAX.
 CC DR PROSITE: PS00197; 2FE2S_FERREDOXIN; 1.
 CC DR PFAM: PF00111; fer2; 1.
 CC DR PFAM: PF00175; Oxidoreductase; Flavoprotein; FMN; NAD;
 CC KW Hypothetical protein; Oxidoreductase; Flavoprotein; FMN; NAD;
 CC KW Iron-sulfur; Electron transport.
 CC FT NP_BIND 6 103 FMN (BY SIMILARITY).
 CC FT NP_BIND 113 226 NAD (BY SIMILARITY).
 CC FT METAL 270 270 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 CC FT METAL 275 275 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 CC FT METAL 278 278 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 CC FT METAL 309 309 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 CC SQ SEQUENCE 321 AA; 35661 MW; 9E85CC68 CRC32;
 CC -----
 CC Query Match 63.3%; Score 31; DB 1; Length 321;
 CC Best Local Similarity 55.6%; Pred. No. 24;
 CC Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC QY 1 YRLAIRLNE 9
 CC |:::|::|
 CC Db 70 YQIAVRLNE 78
 CC -----
 CC RESULT 10
 CC ID PAFL_YEAST STANDARD; PRT; 445 AA.
 CC AC P38351;
 CC DT 01-OCT-1994 (Rel. 30, Created)
 CC DT 01-OCT-1994 (Rel. 30, Last sequence update)
 CC DT 01-NOV-1995 (Rel. 32, Last annotation update)
 CC DE PAFL PROTEIN.
 CC GN PAFL OR YBR279W OR YBR2016.
 CC OS Saccharomyces cerevisiae (Baker's yeast).
 CC OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
 CC OC Saccharomycetales; Saccharomycetes.
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-S288C;
 CC RX MEDLINE: 94378722.
 CC RA HOLMSTROM K., BRANDT T., KALLESOE T.;
 CC "The sequence of a 32,420 bp segment located on the right arm of
 CC chromosome II from Saccharomyces cerevisiae.";
 CC Yeast 10:S47-S62(1994).
 CC [2]
 CC RP IDENTIFICATION.
 CC RA SHI X., WADE P., BURTON Z., JAEHNING J.A.;
 CC Unpublished observations (XXX-1995).
 CC -!- FUNCTION: RNA POLYMERASE II ASSOCIATED PROTEIN IMPORTANT FOR
 CC TRANSCRIPTION.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X76053; CA553642.1; -;
 CC EMBL: Z36148; CA85243.1; -;
 CC PIR: S44541; S44541.
 CC PIR: S39135; S39135.
 CC SGD: L0002621; PAFL.

```

KW Transcription; Nuclear protein.
SQ SEQUENCE 445 AA; 51800 MW; 9259106B CRC32;

Query Match 63.3%; Score 31; DB 1; Length 445;
Best Local Similarity 75.0%; Pred. No. 34;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAIRLNER 10
DB 288 LAIRLNDK 295

RESULT 11
65KD_ZYMO
ID 65KD_ZYMO STANDARD; PRT; 584 AA.
AC P15255;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 65 KD PROTEIN (ORF 1).
OS Zymomonas mobilis.
OC Plasmid p2M2.
OC Bacteria; Proteobacteria; alpha subdivision; Zymomonas group;
OC Zymomonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 10988 / ZM1;
RA MISAWA N., NAKAMURA K.;
RT "The nucleotide sequence of the 2.7 kilobase pair plasmid of Zymomonas
RL J. Biotechnol. 12:63-70(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L24119; AAA70047.1;
DR PFAM; PF00072; response_reg; 1.
DR PFAM; PF00512; signal; 1.
DR PFAM; PF01590; GAF; 1.
KW Sensory transduction; Phosphorylation; Transmembrane; Transferase;
FT DOMAIN 332 572 TRANSMITTER DOMAIN (POTENTIAL).
FT DOMAIN 607 728 RECEIVER DOMAIN (POTENTIAL).
FT TRANSMEM 26 43 POTENTIAL.
FT TRANSMEM 53 76 POTENTIAL.
FT TRANSMEM 83 106 POTENTIAL.
FT MOD_RES 353 353 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 659 659 PHOSPHORYLATION (BY SIMILARITY).
FT MUTAGEN 31 31 A->V: IN ETRI-3; ETHYLENE INSENSITIVITY.
FT MUTAGEN 62 62 I->F: IN ETRI-4; ETHYLENE INSENSITIVITY.
FT MUTAGEN 65 65 C->Y: IN ETRI-1; ETHYLENE INSENSITIVITY.
FT MUTAGEN 102 102 A->T: IN ETRI-2; ETHYLENE INSENSITIVITY.
SQ SEQUENCE 738 AA; 82565 MW; FD593871 CRC32;

Query Match 63.3%; Score 31; DB 1; Length 738;
Best Local Similarity 40.0%; Pred. No. 59;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
DB 668 YQIALRIHEK 677

RESULT 13
YPDQ_BACSU
ID YPDQ_BACSU STANDARD; PRT; 132 AA.
AC P54162;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HYPOTHETICAL 14.7 KD PROTEIN IN BCSA-DEGR INTERGENIC REGION.
GN YPDQ.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-168 / MARBURG;
RA SOROKIN A.V., AZEVEDO V., ZUMSTEIN E., GALLERON N., EHRLICH S.D.,
RR ERROR P.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between
RL the serA and kdg loci cloned in a yeast artificial chromosome.";
RL Microbiology 142:2005-2016(1996).
CC -----
CC -I- SIMILARITY: TO E.FAECALIS CELL WALL ENZYME EBSB.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L77246; AAA96618.1; -
CC EMBL; Z99115; CAB14117.1; -
CC HSSP; P00647; IGOC.
CC SUBTILIST; BG11608; YPDQ.
CC PFAM; PF00075; rna5h; 1.
CC Hypothetical protein.
CC -----
CC Query Match 61.2%; Score 30; DB 1; Length 132;
CC Best Local Similarity 85.7%; Pred. No. 15;
CC Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC -----
QY 4 AIRLNER 10
DB 125 AIRLNER 131
| | | | |
| | | | |

RESULT 14
TRYP_MOUSE STANDARD; PRT; 246 AA.
ID TRY1_MOUSE
AC P07146;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE TRYPSINOGEN PRECURSOR (EC 3.4.21.4).
GN TRY2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-A/J;
RX MEDLINE; 87066713.
RA STEVENSON B.J., HAGENBUCHLE O., WELLAUER P.K.;
RT "Sequence organisation and transcriptional regulation of the mouse
RT elastase II and trypsin genes.";
RL Nucleic Acids Res. 14:8307-8330(1986).
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X04574; CAA28243.1; -
CC EMBL; X04577; CAA28245.1; -
CC PIR; B25528; B25528.
CC HSSP; P00763; LDPO.
CC MGD; MGI:102759; TRY2.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC PFAM; PF00089; trypsin; 1.
CC Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 1 15
FT PROPEP 16 23
FT CHAIN 24 246
FT ACT_SITE 63 63
FT CHARGE RELAY SYSTEM (BY SIMILARITY).

```

```

FT ACT_SITE 107 107
FT ACT_SITE 200 200
FT DISULFID 30 150
FT DISULFID 48 64
FT DISULFID 132 233
FT DISULFID 139 206
FT DISULFID 171 185
FT DISULFID 196 220
FT SITE 134 134
SQ SEQUENCE 246 AA; 26203 MW; BD975983 CRC32;

Query Match 61.2%; Score 30; DB 1; Length 246;
Best Local Similarity 55.6%; Pred. No. 29;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNE 9
DB 67 YRIQVRLGE 75
| | | | |
| | | | |

RESULT 15
YAGI_ECOLI STANDARD; PRT; 252 AA.
ID YAGI_ECOLI
AC P77300;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN PERR-ARGF INTERGENIC REGION.
GN YAGI.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-K12 / MG1655;
RX MEDLINE; 97426617.
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA SCHRAMM S., DUNCAN M., ALLEN E., ARAUJO R., APARICIO A., CHUNG E.,
RA DAVIS K., FEDERSPIEL N., HYMAN R., KALMAN S., KOMP C., KURDI O.,
RA LASHKARI D., LEW H., LIN D., NAMATH A., OEFNER P., ROBERTS D.,
RA DAVIS R.W.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ICLR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF000135; AAC73375.1; -
CC EMBL; U70214; AAB08693.1; -
CC ECOGENE; EG13349; YAGI.
CC PROSITE; PS01051; HTH_ICLR_FAMILY; 1.
CC PFAM; PF01614; ICLR; 1.
CC Hypothetical protein; Transcription regulation; DNA-binding.
FT DNA_BIND 25 45
FT H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 252 AA; 27838 MW; 3D954295 CRC32;

Query Match 61.2%; Score 30; DB 1; Length 252;
Best Local Similarity 60.0%; Pred. No. 30;

```

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YRLAIRLNER 10
 ||| :|| ||
Db 62 YRLGKMLVER 71

Search completed: February 8, 2000, 00:59:50
Job time: 3779 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 13:17:35 ; Search time 209.03 Seconds
(without alignments)
3.317 Million cell updates/sec

Title: US-08-653-294-9
Perfect score: 49
Sequence: 1 YRLAIRLNER 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

SPTREMBL_12:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 36 | 73.5 | 485 | 12 | O39258 |
| 2 | 35 | 71.4 | 2632 | 5 | P90736 equine herp |
| 3 | 33 | 67.3 | 185 | 2 | O55192 caenorhabdi |
| 4 | 33 | 67.3 | 740 | 10 | O82436 cucumis mel |
| 5 | 33 | 67.3 | 741 | 10 | O81122 malus domes |
| 6 | 33 | 67.3 | 1324 | 2 | Q44103 amycolatops |
| 7 | 32 | 65.3 | 123 | 2 | O25306 helicobacte |
| 8 | 32 | 65.3 | 123 | 2 | O92LP7 helicobacte |
| 9 | 32 | 65.3 | 153 | 2 | O31288 buchera ap |
| 10 | 32 | 65.3 | 191 | 12 | Q98543 parametium |
| 11 | 32 | 65.3 | 468 | 10 | O04509 arabidopsis |
| 12 | 32 | 65.3 | 661 | 3 | Q12033 saccharomyc |
| 13 | 32 | 65.3 | 1070 | 2 | Q44070 aeromonas h |
| 14 | 32 | 65.3 | 1154 | 3 | Q01375 neurospora |
| 15 | 32 | 65.3 | 1154 | 3 | Q01379 neurospora |
| 16 | 32 | 65.3 | 2078 | 12 | Q69055 human herpe |
| 17 | 32 | 65.3 | 2282 | 11 | Q61479 mus muscullu |
| 18 | 31 | 63.3 | 144 | 5 | Q27369 trypanosoma |
| 19 | 31 | 63.3 | 377 | 2 | Q9X025 thermotoga |
| 20 | 31 | 63.3 | 395 | 10 | Q9zt98 arabidopsis |

| | | | | | |
|----|----|------|------|----|--------|
| 21 | 31 | 63.3 | 417 | 8 | O20958 |
| 22 | 31 | 63.3 | 418 | 1 | O59284 |
| 23 | 31 | 63.3 | 463 | 1 | O9YAG3 |
| 24 | 31 | 63.3 | 495 | 5 | O16917 |
| 25 | 31 | 63.3 | 517 | 12 | Q98160 |
| 26 | 31 | 63.3 | 538 | 2 | Q9X294 |
| 27 | 31 | 63.3 | 624 | 10 | O81768 |
| 28 | 31 | 63.3 | 717 | 10 | O04325 |
| 29 | 31 | 63.3 | 735 | 10 | O49230 |
| 30 | 31 | 63.3 | 845 | 5 | O01914 |
| 31 | 31 | 63.3 | 1280 | 13 | Q90933 |
| 32 | 31 | 63.3 | 2391 | 10 | Q9XE40 |
| 33 | 30 | 61.2 | 102 | 2 | Q32370 |
| 34 | 30 | 61.2 | 102 | 2 | O46302 |
| 35 | 30 | 61.2 | 107 | 10 | O81521 |
| 36 | 30 | 61.2 | 123 | 10 | O40796 |
| 37 | 30 | 61.2 | 176 | 2 | O83481 |
| 38 | 30 | 61.2 | 201 | 3 | Q12530 |
| 39 | 30 | 61.2 | 336 | 2 | P71615 |
| 40 | 30 | 61.2 | 345 | 2 | O83189 |
| 41 | 30 | 61.2 | 347 | 2 | O85943 |
| 42 | 30 | 61.2 | 381 | 1 | O59512 |
| 43 | 30 | 61.2 | 408 | 2 | P96565 |
| 44 | 30 | 61.2 | 459 | 2 | O59964 |
| 45 | 30 | 61.2 | 517 | 3 | O94556 |

ALIGNMENTS

RESULT 1
O39258 PRELIMINARY; PRT; 485 AA.
ID O39258
AC O39258; 01-JAN-1998 (TREMREL. 05, Created)
DT 01-JAN-1998 (TREMREL. 05, Last sequence update)
DT 01-JAN-1998 (TREMREL. 12, Last annotation update)
DE COUNTERPART OF HSV-1 GENE UL44 AND VZV GENE 14.
GN 16.
OS Equine herpesvirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NS80567;
RX MEDLINE; 91021040.
RA NICOLSON L., ONIONS D.E.;
RT "The nucleotide sequence of the equine herpesvirus 4 gc gene homologues."
RL Virology 179:378-387(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NS80567;
RX MEDLINE; 98264497.
RA TELFORD E.A.R., WATSON M.S., PERRY J., CULLINANE A.A., DAVISON A.J.;
RT "The DNA sequence of equine herpesvirus-4.";
J. Gen. Virol. 79:1197-1203(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-NS80567;
RA TELFORD E.A., WATSON M.S., PERRY J., CULLINANE A.A., DAVISON A.J.;
RT Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF030027; AAC59550.1;
DR PRINTS; PR00668; GLYCOPROTEIN.
SQ SEQUENCE 485 AA; 52539 MW; 8DF52A42 CRC32;

Query Match 73.5% Score 36; DB 12; Length 485;
Best Local Similarity 70.0% Pred. No. 13;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 YRLAIRLNER 10
||| | |||

Db 120 YRLHLNQR 129

RESULT 2

P90736 PRELIMINARY; PRT; 2632 AA.
 ID P90736; 001311;
 AC P90736; 001311;
 DT 01-JAN-1999 (TREMREL. 09, Created)
 DT 01-JAN-1999 (TREMREL. 09, Last sequence update)
 DT 01-JAN-1999 (TREMREL. 09, Last annotation update)
 DE HYPOTHETICAL 305.8 KD PROTEIN B0365.7 IN CHROMOSOME V.
 GN B0365.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA STMS M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 [2]
 RN REVISIONS.
 RC STRAIN-BRISTOL N2;
 RA JONES S.J.M.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 [3]
 RN SEQUENCE FROM N.A.
 RP PERCY C.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: WEAK, TO DYNEIN HEAVY CHAINS.
 DR EMBL; 281028; CAB02695.1;
 DR EMBL; 281096; CAB02695.1; JOINED.
 DR EMBL; 281096; CAB03163.1;
 DR EMBL; 281028; CAB03163.1; JOINED.
 DR WORMPEP; B0365.7; CE07724.
 KW Hypothetical protein.
 SQ SEQUENCE: 2632 AA; 305774 MW; 4F3356EF CRC32;

Query Match 71.4%; Score 35; DB 5; Length 2632;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRLNE 9
 |||:|:|
 Db 557 YRSVRINE 565

RESULT 3

Q55192 PRELIMINARY; PRT; 185 AA.
 ID Q55192;
 AC Q55192;
 DT 01-NOV-1996 (TREMREL. 01, Created)
 DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
 DT 01-JAN-1999 (TREMREL. 09, Last annotation update)
 DE HYPOTHETICAL 20.8 KD PROTEIN
 OS Synchocystis sp. (strain PCC 5803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-PCC6803;
 RA TABATA S.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-PCC6803;
 RX MEDLINE; 96127529.
 RA KANEKO T., TANAKA A., SATO S., KOTANI H., SAZUKA T., MIYAJIMA N.,
 RA SUGIURA M., TABATA S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synchocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
 RT region from map positions 64% to 92% of the genome.";
 RL D.N. Res. 2:153-166(1995).

RESULT 4

Q82436 PRELIMINARY; PRT; 740 AA.
 ID Q82436;
 AC Q82436;
 DT 01-NOV-1998 (TREMREL. 08, Created)
 DT 01-NOV-1998 (TREMREL. 08, Last sequence update)
 DT 01-NOV-1999 (TREMREL. 12, Last annotation update)
 DE PUTATIVE ETHYLENE RECEPTOR.
 GN MEETR1.
 OS Cucumis melo var. reticulatus (netted muskmelon).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eurosids I; Cucurbitales; Cucurbitaceae;
 OC Cucumis.
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=FRUIT;
 RA SATO-NARA K., YUHASHI K., HIGASHI K., HOSoya K., KUBOTA M., EZURA H.;
 RT "Stage- and tissue- specific expression of ETR1 and ERS homolog genes
 during fruit development in melon (Cucumis melo L. Reticulatus).";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF054806; AAC99645.1;
 DR MENDEL; 33025; Cucme; 2316; 33025.
 DR PFAM; PF01590; GAF; 1.
 DR PFAM; PF00072; response_reg; 1.
 DR PFAM; PF00512; signal; 1.
 SQ SEQUENCE 740 AA; 82657 MW; 183F5227 CRC32;

Query Match 67.3%; Score 33; DB 2; Length 185;
 Best Local Similarity 60.0%; Pred. No. 20;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 |||||:|
 Db 35 YRLAIRLQR 44

RESULT 5

Q81122 PRELIMINARY; PRT; 741 AA.
 ID Q81122;
 AC Q81122;
 DT 01-NOV-1998 (TREMREL. 08, Created)
 DT 01-NOV-1998 (TREMREL. 08, Last sequence update)
 DT 01-NOV-1999 (TREMREL. 12, Last annotation update)
 DE ETHYLENE RECEPTOR.

Query Match 67.3%; Score 33; DB 10; Length 740;
 Best Local Similarity 60.0%; Pred. No. 88;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 |||||:|
 Db 671 YELAIRIREK 580

GN ETR1.
 OS Malus domestica (Apple) (Malus sylvestris).
 OC Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eurosids I; Rosales; Rosaceae; Malus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, GRANNY SMITH; TISSUE-RIPENING FRUIT;
 RA LEE S.A., ROSS G.S., GARDNER R.C.;
 RT "An apple (Malus domestica L. Borkh cv Granny Smith) homolog of the
 RT ethylene receptor gene ETR1 (Accession No. AF032448) (PGR98-125).";
 RL Plant Physiol. 117:1126-1126(1998).
 DR EMBL: AF032448; AAC31123.1; -.
 DR MENDEL: 31773; Maldo; 2316:31773.
 DR PFAM: PF01590; GAF; 1.
 DR PFAM: PF00072; response_reg; 1.
 DR PFAM: PF00512; signal; 1.
 DR PFAM: PF00512; signal; 1.
 SQ SEQUENCE 741 AA; 82967 MW; D350B396 CRC32;

Query Match 67.3%; Score 33; DB 10; Length 741;
 Best Local Similarity 50.0%; Pred. No. 88;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 |||:::|
 DB 672 YELAVRIHEK 681

RESULT 6
 Q44103 PRELIMINARY; PRT; 1324 AA.
 AC Q44103;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
 DE PEPTIDE-SYNTHETASE (FRAGMENT).
 GN APS.
 OS Amycolatopsis mediterranei.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Amycolatopsi.
 OC Actinomycetales; Pseudonocardineae; Pseudonocardaceae; Amycolatopsi.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DSM 5908;
 RA MEDLINE; 97449857.
 RA PELZER S., REICHERT W., HUPPERT M., HECKMANN D., WOHLLEBEN W.;
 RT "Cloning and analysis of a peptide synthetase gene of the balhimycin
 RT producer Amycolatopsis mediterranei DSM5908 and development of a gene
 RT disruption/replacement system.";
 RL J. Biotechnol. 56:115-128(1997).
 DR EMBL: X97860; CAA66454.1; -.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 DR PFAM: PF00501; AMP-binding; 1.
 DR PFAM: PF00668; DUF4; 2.
 DR PFAM: PF00550; pp-binding; 1.
 KW Ligase.
 FT NON_TER 1 1
 FT NON_TER 1324 1324
 SQ SEQUENCE 1324 AA; 142666 MW; 2C08588E CRC32;

Query Match 67.3%; Score 33; DB 2; Length 1324;
 Best Local Similarity 70.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 |||:::|
 DB 969 YRVAGRLDER 978

RESULT 7
 Q25306 PRELIMINARY; PRT; 123 AA.
 ID Q25306

AC O25306;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
 DE FLAGELLAR SWITCH PROTEIN (FLIN).
 GN HP0584.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-26695;
 RX MEDLINE; 97394467.
 RA TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,
 RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,
 RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,
 RA LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,
 RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,
 RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,
 RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATTHEY L., WALLIN E.,
 RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,
 RA VENTER J.C.;
 RT "The complete genome sequence of the gastric pathogen Helicobacter
 RT pylori.";
 RL Nature 388:539-547(1997).
 DR EMBL: AE000571; AAD07636.1; -.
 DR TIGR; HP0584;
 DR PRINTS; PR00956; FLGMOTOREFLIN.
 DR Hypothetical protein; Flagella.
 KW SEQUENCE 123 AA; 13942 MW; 498DE0AE CRC32;

Query Match 65.3%; Score 32; DB 2; Length 123;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAIRLINE 9
 |||:::|
 DB 102 LAIRLINE 108

RESULT 8
 Q92LP7 PRELIMINARY; PRT; 123 AA.
 ID Q92LP7
 AC Q92LP7; 1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
 DE PUTATIVE FLAGELLAR MOTOR SWITCH PROTEIN.
 GN FLIN.
 OS Helicobacter pylori J99.
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-J99;
 RX MEDLINE; 99120557.
 RA ALM R.A., LING L.-S.L., MOIR D.T., KING B.L., BROWN E.D., DOIG P.C.,
 RA SMITH D.R., NOONAN B., GUILD B.C., DEJONGE B.D., CARMEL G.,
 RA TUMMINO P.J., CARUSO A., URIA-NICKELSEN M., MILLS D.M., IVES C.,
 RA GIBSON R., MERBERG D., MILLS S.D., JIANG Q., TAYLOR D.E., VOVIS G.F.,
 RA TRUST T.J.;
 RT "Genomic-sequence comparison of two unrelated isolates of the human
 RT gastric pathogen Helicobacter pylori.";
 RL Nature 397:176-180(1999).
 DR EMBL: AE001485; AAD06097.1; -.
 SQ SEQUENCE 123 AA; 13966 MW; 9406C37C CRC32;

Query Match 65.3%; Score 32; DB 2; Length 123;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LAIRLINE 9
 DB 102 LAIRLINE 108

RESULT 9

O31288 PRELIMINARY; PRT; 153 AA.
 AC O31288;
 DT 01-JAN-1998 (TRENBLrel. 05, Created)
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
 DE HSPA PROTEIN.
 GN HSPA.
 OS Buchnera aphidicola.
 OG Plasmid pBtS1.
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 97386415.
 RA VAN HAM R.C.H.J., MOYA A., LATORRE A.;
 RT "Putative evolutionary origin of plasmids carrying the genes involved
 RT in leucine biosynthesis in Buchnera aphidicola (endosymbiont of
 RT aphids).";
 RL J. Bacteriol. 179:4768-4777(1997).
 DR EMBL: Y11986; CAA72698.1; -;
 KW Plasmid.
 SQ SEQUENCE 153 AA; 17953 MW; 66A38676 CRC32;

Query Match 65.3%; Score 32; DB 2; Length 153;

Best Local Similarity 60.0%; Pred. No. 27;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 DB 109 FSLSFRLNER 118

RESULT 10

Q98543 PRELIMINARY; PRT; 191 AA.
 AC Q98543;
 DT 01-FEB-1997 (TRENBLrel. 02, Created)
 DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
 DE GENOME, PARTIAL SEQUENCE.
 GN A493L.
 OS Parametrium bursaria chlorella virus 1 (PBCV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phycodnavirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 96400190.
 RA KUTISH G.F., LI Y., LU Z., FURUTA M., ROCK D.L., VAN ETTEN J.L.;
 RT "Analysis of 76 kb of the chlorella virus PBCV-1 330-kb genome: map
 RT positions 182 to 258.";
 RL Virology 223:303-317(1996).
 DR EMBL: U42580; AAC96860.1; -;
 SQ SEQUENCE 191 AA; 22651 MW; E4547C83 CRC32;

Query Match 65.3%; Score 32; DB 12; Length 191;

Best Local Similarity 66.7%; Pred. No. 34;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLINE 9
 DB 74 YRLVRENE 82

RESULT 11

O04509 PRELIMINARY; PRT; 468 AA.
 ID O04509

AC O04509;
 DT 01-JUL-1997 (TRENBLrel. 04, Created)
 DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE HYPOTHETICAL 52.8 KD PROTEIN F21M12.26 IN CHROMOSOME 1.
 GN F21M12.26.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 OC Arabidopsis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA VISITSRAIA V.S., OSBORNE B.I., TORIUMI M., YU G., OJI O., SHEN Y.K.,
 RA ARAJO R., AU M., BUEHLER E., CONWAY A.B., CONWAY A.R., DEWAR K.,
 RA FENG J., KIM C., KURTZ D., LI Y., SHINN P., SUN H., DAVIS R.W.,
 RA ECKER J.R., FEDERSPIEL N.A., THEOLOGIS A.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC000132; AAB60740.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 468 AA; 52790 MW; E9129511 CRC32;

Query Match 65.3%; Score 32; DB 10; Length 468;

Best Local Similarity 60.0%; Pred. No. 89;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 DB 132 YOLGIRVRER 141

RESULT 12

Q12033 PRELIMINARY; PRT; 661 AA.
 AC Q12033;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE HYPOTHETICAL 75.9 KD PROTEIN IN VP1-SNF2 INTERGENIC REGION.
 GN YOR275C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHERET G., SOR F.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RX MEDLINE; 97051594.
 RA CHERET G., BERNARDI A., SOR F.;
 RT "DNA sequence analysis of the VP1-SNF2 region on chromosome XV of
 RT Saccharomyces cerevisiae.";
 RL Yeast 12:1059-1064(1996).
 CC -!- SIMILARITY: SOME, TO C.ELEGANS R10E12.1.
 DR EMBL: 275183; CAA99500.1; -;
 DR EMBL: X89633; CAA61781.1; -;
 DR SGD: S0005801; YOR275C.
 KW Hypothetical protein.
 SQ SEQUENCE 661 AA; 75947 MW; 266DA004 CRC32;

Query Match 65.3%; Score 32; DB 3; Length 661;

Best Local Similarity 50.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 DB 241 YRALSFNK 250

RESULT 13

Q44070 PRELIMINARY; PRT; 1070 AA.
 AC Q44070; 1070 AA.
 DT 01-NOV-1996 (TREMELREL. 01, Created)
 DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)
 DT 01-NOV-1998 (TREMELREL. 08, Last annotation update)
 DE NUCLEASE.
 GN NUCH.
 OS Aeromonas hydrophila.
 OC Bacteria; Proteobacteria; gamma subdivision; Aeromonas group;
 OC Aeromonas.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JMP636.
 RX MEDLINE; 96272269.
 RA DODD H.N., PEMBERTON J.M.;
 RT "Cloning, sequencing, and characterization of the nuch gene encoding
 RT an extracellular nuclease from Aeromonas hydrophila JMP636.";
 RL J. Bacteriol. 178:3926-3933(1996).
 DR EMBL; L76304; AAB39273.1;
 KW Nuclease.
 SQ SEQUENCE 1070 AA; 113719 MW; EC9538BF CRC32;

Query Match 65.3%; Score 32; DB 2; Length 1070;
 Best Local Similarity 60.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNER 10
 ||||: ||
 DB 54 YRLALYANDR 63

RESULT 14

Q01375 PRELIMINARY; PRT; 1154 AA.
 AC Q01375;
 DT 01-NOV-1996 (TREMELREL. 01, Created)
 DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)
 DT 01-NOV-1998 (TREMELREL. 08, Last annotation update)
 DE HYPOTHETICAL 130.4 KD PROTEIN.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Euascomycetes; Pyrenomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-J1518;
 RX MEDLINE; 94203179.
 RA CAMBARERI E.B., HELBER J., KINSEY J.A.;
 RT "Tad1-1, an active LINE-like element of Neurospora crassa.";
 RL Mol. Gen. Genet. 242:658-665(1994).
 DR EMBL; L25662; AAA21781.1;
 DR PFAM; PF00078; rvt; 1.
 KW Hypothetical protein.
 FT DOMAIN 1019 1022 POLY-LYS.
 FT DOMAIN 1029 1034 POLY-GLU.
 SQ SEQUENCE 1154 AA; 130398 MW; DF0BA680 CRC32;

Query Match 65.3%; Score 32; DB 3; Length 1154;
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNE 9
 ||||: ||
 DB 1136 YRLAVELEE 1144

RESULT 15

Q01379 PRELIMINARY; PRT; 1154 AA.
 ID Q01379

AC Q01379;
 DT 01-NOV-1996 (TREMELREL. 01, Created)
 DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)
 DT 01-NOV-1998 (TREMELREL. 08, Last annotation update)
 DE HYPOTHETICAL 130.5 KD PROTEIN.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Euascomycetes; Pyrenomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-J1518;
 RX MEDLINE; 94203179.
 RA CAMBARERI E.B., HELBER J., KINSEY J.A.;
 RT "Tad1-1, an active LINE-like element of Neurospora crassa.";
 RL Mol. Gen. Genet. 242:658-665(1994).
 DR EMBL; L25663; AAA21792.1;
 DR PFAM; PF00078; rvt; 1.
 KW Hypothetical protein.
 FT DOMAIN 1019 1022 POLY-LYS.
 FT DOMAIN 1029 1034 POLY-GLU.
 SQ SEQUENCE 1154 AA; 130470 MW; 7FBE8EAF CRC32;

Query Match 65.3%; Score 32; DB 3; Length 1154;
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 YRLAIRLNE 9
 ||||: ||
 DB 1136 YRLAVELEE 1144

Search completed: February 8, 2000, 13:17:37
 Job time: 32486 sec

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-9 to: GenEmbl:* out_format : pfs

Date: Feb 8, 2000 4:37 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+P2n.model -DEV=xlp
-Q/Cgml_1/USPFO_spool/US08653294/runat_04022000_160701_15779/app_query.fasta.1
-DB=GenEmbl -QWMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.500
-MINMATCH=0.100 -LOOPEXT=0.000 -KEYWORD=4.500
-FGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -ALIGN=15 -MODE=LOCAL
-OUTWMT=pfs -NORMAL=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294
-NCPU=6 -ICPU=3 -NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-9

Query length: 10

Database: GenEmbl.*

Database sequences: 821193

Database length: -1518192014

Search time (sec): 11370.480000

score_list:

| Strd Orig | ZScore | Escore | Len | Documentation |
|-------------------|---------|--------|---------|---------------------------------------|
| gb.pr2:AC002465 | + 42.00 | 104.99 | 796.71 | 155881 AC002465 Human BAC clone RG3 |
| gb.pr5:AC016116 | + 39.00 | 101.04 | 1.3e+03 | 60193 AC016116 Homo sapiens clone R |
| gb.pr3:AC005365 | + 39.00 | 97.85 | 2.0e+03 | 86130 AC005365 Homo sapiens chromos |
| gb.htg1:HS1141E20 | - 39.00 | 96.71 | 2.3e+03 | 97906 AL109912 Homo sapiens chromos |
| gb.pr3:AC004606 | - 39.00 | 96.56 | 2.3e+03 | 99512 AC004606 Homo sapiens chromos |
| gb.pr3:HS1111D26 | - 39.00 | 93.82 | 3.3e+03 | 135305 AL080317 Human DNA sequence |
| gb.pr3:CN5018P2 | - 39.00 | 90.96 | 4.8e+03 | 186531 AL109985 Human chromosome 14 |
| gb.ov:D64134 | - 38.00 | 129.36 | 34.97 | 1579 D64134 Novt mRNA for annexin v |
| gb.htg4:AC010886 | + 38.00 | 86.64 | 8.4e+03 | 190094 AC010886 Homo sapiens chrom |
| gb.htg3:AC009685 | + 38.00 | 85.75 | 9.4e+03 | 210031 AC009685 Homo sapiens chrom |
| gb.in1:CEY758A | - 38.00 | 82.62 | 1.4e+04 | 298406 AL033514 Caenorhabditis eleg |
| gb.htg1:CEY758B | - 38.00 | 81.55 | 1.6e+04 | 336638 AL022286 Caenorhabditis eleg |
| gb.sts:G46114 | - 37.00 | 135.75 | 15.43 | 485 G46114 Z6550.1 Zebrafish AB Dan |
| gb.ro:AB012933 | - 37.00 | 121.29 | 98.54 | 2454 AB012933 Rattus norvegicus mRN |
| gb.p11:YGCCOT1A | + 37.00 | 120.52 | 108.76 | 2675 M82522 Saccharomyces cerevisia |
| gb.htg7:AC017159 | + 37.00 | 97.63 | 2.0e+03 | 34827 AC017159 Drosophila melanog |
| gb.in2:AC004287 | + 37.00 | 91.30 | 4.6e+03 | 70841 AC004287 Drosophila melanog |
| gb.in2:AC004713 | - 37.00 | 91.12 | 4.7e+03 | 72327 AC004713 Drosophila melanog |
| gb.htg7:AC017187 | - 37.00 | 88.82 | 6.3e+03 | 93588 AC017187 Drosophila melanog |
| gb.htg7:AC017024 | + 37.00 | 88.76 | 6.4e+03 | 94183 AC017024 Homo sapiens clone R |
| gb.htg4:AC010063 | + 37.00 | 87.36 | 7.6e+03 | 110157 AC010063 Drosophila melanog |
| gb.htg4:AC009659 | + 37.00 | 82.30 | 1.5e+04 | 194446 AC009659 Homo sapiens chrom |
| gb.pat:AG21044 | + 36.00 | 121.18 | 99.88 | 1560 A21044 Equine herpesvirus type |
| gb.vi:ISGSP13 | + 36.00 | 115.72 | 201.37 | 2880 M58031 Equine herpesvirus 4 g |
| gb.p12:AF064563 | + 36.00 | 112.88 | 289.85 | 3960 AF064563 Hordeum vulgare culti |
| gb.p12:AF064563 | + 36.00 | 106.55 | 652.82 | 8094 D37799 Bacillus subtilis genes |
| gb.b2:AC001206 | + 36.00 | 102.97 | 1.0e+03 | 12023 A46241 Caenorhabditis elegans |
| gb.in1:CEC38D4 | - 36.00 | 94.03 | 3.3e+03 | 32777 Z46241 Caenorhabditis elegans |
| gb.b2:AF012285 | + 36.00 | 90.84 | 4.9e+03 | 46864 AF012285 Bacillus subtilis mc |
| gb.pr2:AF000266 | + 36.00 | 87.92 | 7.1e+03 | 65058 AP000266 Homo sapiens genom |
| gb.pr2:AF000034 | - 36.00 | 84.08 | 1.2e+04 | 100000 AP000034 Homo sapiens genom |
| gb.pr2:AF000101 | + 36.00 | 84.08 | 1.2e+04 | 100000 AP000101 Homo sapiens genom |
| gb.pr2:AF000177 | + 36.00 | 84.08 | 1.2e+04 | 100000 AP000177 Homo sapiens genom |
| gb.pr2:AC015845_2 | - 36.00 | 83.23 | 1.3e+04 | 110000 Continuation (3 of 6) of AC0 |
| gb.p11:AF000815 | - 36.00 | 80.93 | 1.7e+04 | 142418 AP000815 Oriza sativa genom |
| gb.htg5:AC011687 | - 36.00 | 80.91 | 1.7e+04 | 142688 AC011687 Homo sapiens clone |
| gb.vi:AF030027 | - 36.00 | 80.73 | 1.8e+04 | 145597 AF030027 Equine herpesvirus |
| gb.pr2:CN501DIF | - 36.00 | 80.14 | 1.9e+04 | 155618 AL132716 Human chromosome 14 |
| gb.htg2:CN501DUR | - 36.00 | 79.78 | 2.0e+04 | 162014 AL133316 Homo sapiens chrom |
| gb.htg2:AC007996 | + 36.00 | 79.23 | 2.2e+04 | 172329 AC007996 Homo sapiens clone |
| gb.htg2:AC007638 | + 36.00 | 78.94 | 2.2e+04 | 178053 AC007638 Homo sapiens chrom |

gb.htg3:AC009485 + 36.00 78.33 2.4e+04 190706 | AC009485 Homo sapiens clo
gb_bai:BSDB0008 + 36.00 77.54 2.7e+04 208230 | Z99111 Bacillus subtilis
gb_bai:RPX01 - 36.00 74.82 3.8e+04 282610 | AJ235270 Rickettsia prow

seq_name: gb_pr2:AC002465

seq_documentation_block:

LOCUS AC002465 155881 bp DNA PRI 20-AUG-1997
DEFINITION Human BAC clone RG343P13 from 7q31, complete sequence.

ACCESSION AC002465

VERSION AC002465.1 GI:2337862

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 155881)

AUTHORS Connell,M and Cordes,M.

TITLE The sequence of H. sapiens BAC clone RG343P13

JOURNAL Unpublished (1997)

2 (bases 1 to 155881)

AUTHORS Waterston,R.

REFERENCE Direct Submission

Submitted (20-AUG-1997) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

SUBMITTED BY:

Genome Sequencing Center

Department of Genetics

Washington University

St. Louis MO 63108, USA

http://genome.wustl.edu/gsc

mailto:sapiens@watson.wustl.edu

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7> or <mailto:regreen@nhgri.nih.gov>

SOURCE INFORMATION:

This clone is from a release of the human BAC library. The library contains cloned DNA from human sperm. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); Kim et al., Genomics 34:213-8 (1996). The clone is available from Research Genetics, Inc. (<http://www.resgen.com>).

VECTOR: pBelOBAC11

Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The actual start of this clone is at base position 1 of RG343P13; actual end is at 155881 of RG343P13. The orientation of this clone is unknown.

This clone contains STS's SWSS847 (NID:g1916380), SWSS846 (NID:g1916379).

Location/Qualifiers

1. 155881

/organism="Homo sapiens"

/db_xref="taxon:9606"

FEATURES

source

```

/chromosome="7"
/clone="RG343P13"
/clone_lib="CITB-HS-A"
/map="7q31"
complement(1528..1820)
/rpt_family="ALU"
7349..7634
/rpt_family="ALU"
7986..8279
/rpt_family="ALU"
complement(9350..9639)
/rpt_family="ALU"
complement(12067..12358)
/rpt_family="ALU"
13275..13399
/rpt_family="ALU"
complement(14341..14480)
/rpt_family="ALU"
15092..15381
/rpt_family="ALU"
complement(17529..17751)
/rpt_family="ALU"
complement(17893..18003)
/notes="match to human EST T99653 (NID:g749390) ye67a02.r1"
complement(17893..18004)
/notes="match to human EST W78848 (NID:g1389395)
zh5ig03.r1"
complement(<17894..43271)
/genes="WNT2"
complement(join(<17894..18158,35353..35630,40849..41075,
43189..43271))
/genes="WNT2"
/notes="match to human mRNA X07876 (NID:g33970) and WNT2
protein P09544 (PID:g139750)"
/codon_start=1
/product="secreted growth factor"
/db_xref="GI:2337863"
/translation="MNPALGIGLWLPILLITLTPVNSWNTWTRATGSSRVWCNDV
RGLVSRQICRRHPDVAISQVAEWTAECQHFQHRWNCNTLDRHSLGRLVLL
RSSRESAFVVAISAGVYFAITRACSQGEVKSCDPKPKSAKDSKGFDFWGCSDN
IDYGIKFAFVAKERKGDARALMNLHNNRAGRKVRKFLKQCKCKGVSGSCITLR
TCWLAMADFRTKGYLWRKYNGALQVVMQDGTGFTVANERFKKPTKNDLVYFENSPD
YCIIDREA"
complement(19783..20079)
/rpt_family="ALU"
complement(20479..20766)
/rpt_family="ALU"
20803..21090
/rpt_family="ALU"
complement(22097..22377)
/rpt_family="ALU"
22508..22606
/rpt_family="L1"
complement(24726..25017)
/rpt_family="ALU"
complement(25953..26002)
/rpt_family="ALU"
complement(26028..26117)
/rpt_family="ALU"
complement(26429..26477)
/rpt_family="L1"
complement(26819..26845)
/rpt_family="L1"
29184..29475
/rpt_family="ALU"
complement(30218..30369)
/rpt_family="ALU"
30565..30874
/rpt_family="ALU"
31884..31921
/rpt_family="L1"
complement(32427..32453)

repeat_region /rpt_family="L1"
complement(32454..32587)
repeat_region /rpt_family="ALU"
complement(32617..32681)
misc_feature 35367..35529
/genes="WNT2"
/notes="match to human EST T29432 (NID:g611530)"
complement(38854..39080)
/rpt_family="ALU"
43070..45031
/notes="CpG island (%GC=64.4, o/e=0.70, #CpGs=157)"
47242..47360
/rpt_family="ALU"
complement(48223..48514)
/rpt_family="ALU"
48717..48995
/rpt_family="ALU"
complement(49005..49032)
/rpt_family="L1"
51231..51772
/rpt_family="L1"
52336..52494
/rpt_family="L1"
52528..52987
/rpt_family="L1"
54271..54366
/rpt_family="L1"
55051..55074
/rpt_family="L1"
55654..55791
/rpt_family="L1"
55902..55935
/rpt_family="L1"
56087..56189
/rpt_family="L1"
complement(56256..56548)
/rpt_family="ALU"
56655..56695
/rpt_family="L1"
56730..56802
/rpt_family="L1"
56818..56938
/rpt_family="ALU"
57144..57227
/rpt_family="L1"
complement(59426..59717)
/rpt_family="ALU"
complement(60815..61106)
/rpt_family="ALU"
complement(61569..61869)
/rpt_family="ALU"
complement(63011..67613)
/rpt_family="L1"
63980..64398
/rpt_family="L1"
67964..68015
/rpt_family="MER"
68059..69083
/rpt_family="MER"
72312..72350

alignment_scores:
Quality: 42.00 Length: 10
Ratio: 4.200 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 70.000

alignment_block:
US-08-653-294-9 x AC002465
Align seg 1/1 to: AC002465 from: 1 to: 155881
1 TyrArgLeuAlaIleArgLeuAsnGluArg 10

```

```
|||||.....:|||||:|||||
15573 TATAGACITTCGTGAGAGTCATGAGAGA 15602

seq_name: gb_htg5:AC016116
seq_documentation_block:
LOCUS AC016116 60193 bp DNA HTG 20-NOV-1999
DEFINITION Homo sapiens clone RP11-27G6, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC016116
VERSION AC016116.1 GI:6456198
KEYWORDS HTG; HTGS_PHASEO.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE 1 (bases 1 to 60193)
JOURNAL Homo sapiens, clone RP11-27G6
REFERENCE
AUTHORS 2 (bases 1 to 60193)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Bouhgafter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donellan,L., Doyle,M.,
Ferrelira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Project Information
Center project name: L4694
Center clone name: 27_G_6
-----
* NOTE: This record contains 77 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 774: contig of 774 bp in length
* gap of unknown length
* 775 1558: contig of 784 bp in length
* gap of unknown length
* 1559 2319: contig of 761 bp in length
* gap of unknown length
* 2320 3094: contig of 775 bp in length
* gap of unknown length
* 3095 3865: contig of 771 bp in length
* gap of unknown length
* 3866 4655: contig of 790 bp in length
* gap of unknown length
* 4656 5455: contig of 800 bp in length
* gap of unknown length
* 5456 6241: contig of 786 bp in length
* gap of unknown length
* 6242 7011: contig of 770 bp in length
* gap of unknown length
*
7790: contig of 779 bp in length
gap of unknown length
8532: contig of 742 bp in length
gap of unknown length
9308: contig of 776 bp in length
gap of unknown length
10078: contig of 770 bp in length
gap of unknown length
10849: contig of 771 bp in length
gap of unknown length
11640: contig of 791 bp in length
gap of unknown length
12428: contig of 788 bp in length
gap of unknown length
13213: contig of 785 bp in length
gap of unknown length
14001: contig of 788 bp in length
gap of unknown length
14789: contig of 788 bp in length
gap of unknown length
15545: contig of 756 bp in length
gap of unknown length
16338: contig of 793 bp in length
gap of unknown length
17118: contig of 780 bp in length
gap of unknown length
17903: contig of 785 bp in length
gap of unknown length
18653: contig of 750 bp in length
gap of unknown length
19452: contig of 799 bp in length
gap of unknown length
20234: contig of 782 bp in length
gap of unknown length
21028: contig of 794 bp in length
gap of unknown length
21805: contig of 777 bp in length
gap of unknown length
22572: contig of 767 bp in length
gap of unknown length
23364: contig of 792 bp in length
gap of unknown length
24143: contig of 779 bp in length
gap of unknown length
24922: contig of 779 bp in length
gap of unknown length
25678: contig of 756 bp in length
gap of unknown length
26459: contig of 781 bp in length
gap of unknown length
27256: contig of 797 bp in length
gap of unknown length
28051: contig of 795 bp in length
gap of unknown length
28851: contig of 800 bp in length
gap of unknown length
29642: contig of 791 bp in length
gap of unknown length
30437: contig of 795 bp in length
gap of unknown length
31224: contig of 787 bp in length
gap of unknown length
31995: contig of 771 bp in length
gap of unknown length
32774: contig of 779 bp in length
gap of unknown length
33540: contig of 766 bp in length
gap of unknown length
34321: contig of 781 bp in length
gap of unknown length
35131: contig of 810 bp in length
gap of unknown length
35933: contig of 802 bp in length
gap of unknown length
7012
7791
8533
9309
10079
10850
11641
12429
13214
14002
14790
15546
16339
17119
17904
18654
19453
20235
21029
21806
22573
23365
24144
24923
25679
26460
27257
28052
28852
29643
30438
31225
31996
32775
33541
34322
35132
```

```
* * *
* 35934 gap of unknown length
* 36692: contig of 759 bp in length
* 36693 gap of unknown length
* 37486: contig of 794 bp in length
* 37487 gap of unknown length
* 38265: contig of 779 bp in length
* 38266 gap of unknown length
* 39051: contig of 786 bp in length
* 39052 gap of unknown length
* 39837: contig of 786 bp in length
* 39838 gap of unknown length
* 40622: contig of 785 bp in length
* 40623 gap of unknown length
* 41392: contig of 770 bp in length
* 41393 gap of unknown length
* 42155: contig of 763 bp in length
* 42156 gap of unknown length
* 42942: contig of 787 bp in length
* 42943 gap of unknown length
* 43718: contig of 776 bp in length
* 43719 gap of unknown length
* 44506: contig of 788 bp in length
* 44507 gap of unknown length
* 45284: contig of 778 bp in length
* 45285 gap of unknown length
* 46062: contig of 778 bp in length
* 46063 gap of unknown length
* 46845: contig of 783 bp in length
* 46846 gap of unknown length
* 47635: contig of 790 bp in length
* 47636 gap of unknown length
* 48424: contig of 789 bp in length
* 48425 gap of unknown length
* 49194: contig of 770 bp in length
* 49195 gap of unknown length
* 49995: contig of 801 bp in length
* 49996 gap of unknown length
* 50764: contig of 769 bp in length
* 50765 gap of unknown length
* 51536: contig of 772 bp in length
* 51537 gap of unknown length
* 52324: contig of 788 bp in length
* 52325 gap of unknown length
* 53118: contig of 794 bp in length
* 53119 gap of unknown length
* 53897: contig of 779 bp in length
* 53898 gap of unknown length
* 54685: contig of 788 bp in length
* 54686 gap of unknown length
* 55463: contig of 778 bp in length
* 55464 gap of unknown length
* 56231: contig of 768 bp in length
* 56232 gap of unknown length
* 57030: contig of 799 bp in length
* 57031 gap of unknown length
* 57826: contig of 796 bp in length
* 57827 gap of unknown length
* 58618: contig of 792 bp in length
* 58619 gap of unknown length
* 59407: contig of 789 bp in length
```

```
alignment_scores:
  Quality: 39.00 Length: 9
  Ratio: 4.333 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 88.889

alignment_block:
US-08-653-294-9 x AC016116/rev ..
Align seg 1/1 to reverse of: AC016116 from: 1 to: 60193
2 ArgLeuAlaIleArgLeuAsnGluArg 10
|||||
```

50768 AGCTAGCAATTCGACTTAACGAGAAA 50742

seq_name: gb_pr3:AC005365

seq_documentation_block:

LOCUS AC005365 86130 bp DNA PRI 01-AUG-1998
DEFINITION Homo sapiens chromosome 16, P1 clone 79-2A (LANL), complete
sequence.

ACCESSION AC005365

VERSION AC005365.1 GI:3367509

KEYWORDS HTG.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 86130)

AUTHORS

Ricke, D.O., Bruce, D., Mundt, M., Doggett, N., Munk, C., Saunders, E.,
Robinson, D., Jones, M., Buckingham, J., Chasteen, L., Thompson, S.,
Goodwin, L., Bryant, J., Tesmer, J., Meincke, L., Longmire, J.,
White, S., Ueng, S., Tatum, O., Campbell, C., Fawcett, J., Maitble, M.,
Misra, M. and Deaven, L.

TITLE Sequencing of Human Chromosome 16

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 86130)

AUTHORS

Ricke, D.O.

TITLE

Large Scale Sequence Analysis and Annotation with the Sequence

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 86130)

AUTHORS

Ricke, D.O., Bruce, D., Mundt, M., Doggett, N., Munk, C., Saunders, E.,
Robinson, D., Jones, M., Buckingham, J., Chasteen, L., Thompson, S.,
Goodwin, L., Bryant, J., Tesmer, J., Meincke, L., Longmire, J.,
White, S., Ueng, S., Tatum, O., Campbell, C., Fawcett, J., Maitble, M.,
Misra, M. and Deaven, L.

TITLE Direct Submission

JOURNAL

Submitted (01-AUG-1998) Center for Human Genome Studies, DOE Joint
Genome Institute, Los Alamos National Laboratory, MS M888, Los
Alamos, NM 87545, USA

FEATURES

source

1. 86130
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="79-2A"

misc_feature

2000..2142

/note="GRAIL 2 excellent exon, frame 2"

repeat_region

2185..2497

/rpt_family="Alu"

repeat_region

7853..8448

/rpt_family="Alu"

repeat_region

9717..10001

/rpt_family="Alu"

repeat_region

10542..10563

/note="(T)22"

/rpt_type=tandem

repeat_region

complement(10546..10672)

/rpt_family="Alu"

misc_feature

11817..12730

/note="GRAIL 2 excellent exon, frame 0"

misc_feature

14804..14902

/note="GRAIL 2 excellent exon, frame 2"

misc_feature

15869..16008

/note="GRAIL 2 excellent exon, frame 1"

repeat_region

16275..16548

/rpt_family="Alu"

repeat_region

17124..17355

/rpt_family="MER33"

repeat_region

complement(18448..18747)

/rpt_family="Alu"

repeat_region

19083..19373

/rpt_family="Alu"

repeat_region

19357..19380

/note="(A)24"

| | |
|---|---|
| <p>seq_name: gb_htgl:HS1141E20</p> <p>seq_documentation_block: LOCUS HS1141E20 97906 bp DNA HTG 23-NOV-1999 DEFINITION Homo sapiens chromosome 6 clone RP5-1141E20, *** SEQUENCING IN PROGRESS ***, in unordered pieces. ACCESSION ALI09912 KEYWORDS HTG; HTGS_PHASE1. SOURCE human.</p> <p>ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.</p> <p>REFERENCE AUTHORS Sims.S. TITLE Direct Submission JOURNAL Submitted (10-SEP-1999) Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@anger.ac.uk Clone requests: clonerequest@anger.ac.uk On Sep 12, 1999 this sequence version replaced gi:5777438. IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, phage etc. Order of segments is not known; 800 n's separate segments. Unfinished: djl1141E20 Contig.ID: 00340 acc=ALI09912 Length: 64235 bp Unfinished: djl1141E20 Contig.ID: 00618 acc=ALI09912 Length: 18869 bp Unfinished: djl1141E20 Contig.ID: 00679 acc=ALI09912 Length: 13202 bp. * NOTE: This is a 'working draft' sequence. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved.</p> | <p>TITLE Location/Qualifiers source 1. .97906 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="6" /clone="RP5-1141E20" /clone-lib="RPCI-5"</p> <p>BASE COUNT 30195 a 18298 c 18016 g 29795 t 1602 others</p> <p>ORIGIN</p> <p>alignment_scores: Quality: 39.00 Length: 10 Ratio: 4.333 Gaps: 0 Percent Similarity: 90.000 Percent Identity: 80.000</p> <p>alignment_block: US-08-653-294-9 x HS1141E20/rev .. Align seg 1/1 to reverse of: HS1141E20 from: 1 to: 97906 1 TyrArgLeuAlaLleArgLeuAnGLuarG 10 48886 TACAGATTGTTCATCAGATTGAATCCAGA 48857</p> <p>seq_name: gb_pr3:AC004606</p> <p>seq_documentation_block: LOCUS AC004606 99512 bp DNA PRI 01-MAY-1998 DEFINITION Homo sapiens chromosome 17, clone hRPC.1043_H_15, complete sequence. ACCESSION AC004606 VERSION AC004606.1 GI:3097870 KEYWORDS HTG. SOURCE human. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.</p> |
|---|---|

[illegible]


```
repeat_region /rpt_family="LI"
repeat_region 8108..8301
repeat_region /rpt_family="THE1C"
repeat_region 8304..8559
repeat_region /rpt_family="LIP2"
repeat_region complement(8572..10975)
repeat_region /rpt_family="LIM4A"
repeat_region complement(11007..11074)
repeat_region /rpt_family="(TA)n"
repeat_region complement(11076..11162)
repeat_region /rpt_family="LIM4A"
repeat_region complement(11172..11276)
repeat_region /rpt_family="AluSg/x"
repeat_region 11282..11677
repeat_region /rpt_family="THE1C"
repeat_region complement(11678..11766)
repeat_region /rpt_family="Alu"
repeat_region complement(11797..11843)
repeat_region /rpt_family="(CA)n"
repeat_region complement(11845..12028)
repeat_region /rpt_family="LIM4A"
repeat_region 12027..12729
repeat_region /rpt_family="LI"
repeat_region 12715..13454
repeat_region /rpt_family="LIM4B"
repeat_region 13452..13666
repeat_region /rpt_family="LI"
repeat_region 13690..14077
repeat_region /rpt_family="MSTA"
repeat_region complement(14086..14412)
repeat_region /rpt_family="LIME1"
repeat_region complement(14391..14636)
repeat_region /rpt_family="LIMC/D"
repeat_region complement(14632..14760)
repeat_region /rpt_family="LIM4"
repeat_region 15025..15291
repeat_region /rpt_family="LIP4B"
repeat_region 15292..15334
repeat_region /rpt_family="AT-rich"
repeat_region 15477..15568
repeat_region /rpt_family="LIM4"
repeat_region 16441..17498
repeat_region /rpt_family="LIME3A"
repeat_region complement(18781..18916)
repeat_region /rpt_family="LIP412"
repeat_region 18907..19237
repeat_region /rpt_family="LIP412"
repeat_region complement(19315..19496)
repeat_region /rpt_family="MIR"
repeat_region 21444..21504
repeat_region /rpt_family="L2"
repeat_region complement(21505..21870)
repeat_region /rpt_family="MLT1A1"
repeat_region 21942..22030
repeat_region /rpt_family="MIR"
repeat_region complement(22031..22121)
repeat_region /rpt_family="LIP4G"
repeat_region 22119..23393
repeat_region /rpt_family="LIP4G"
repeat_region complement(23440..23801)
repeat_region /rpt_family="LIP41"
repeat_region complement(23808..23868)
repeat_region /rpt_family="(TA)n"
repeat_region complement(23882..24177)
repeat_region /rpt_family="AluSg"
repeat_region complement(24187..24823)
repeat_region /rpt_family="LIP41"
repeat_region complement(24826..25342)
repeat_region /rpt_family="LI"
repeat_region 25355..25488
repeat_region /rpt_family="MIR"
repeat_region 25948..26024
repeat_region /rpt_family="MLT1D"
```

```
repeat_region complement(27114..27413)
repeat_region /rpt_family="AluSc"
repeat_region complement(27414..27606)
repeat_region /rpt_family="MLT2A"
repeat_region complement(27905..27992)
repeat_region /rpt_family="LIME1"
repeat_region 28057..28462
repeat_region /rpt_family="MSTA"
repeat_region complement(28466..28816)
repeat_region /rpt_family="MLT1A2"
repeat_region complement(28827..29361)
repeat_region /rpt_family="LIME1"
repeat_region complement(29387..29618)
repeat_region /rpt_family="LIMB8"
repeat_region 30005..30032
repeat_region /rpt_family="(GA)n"
repeat_region 30643..30889
repeat_region /rpt_family="LIM4"
repeat_region 30890..31792
repeat_region /rpt_family="LIM4"
repeat_region 31804..31870
repeat_region /rpt_family="AT-rich"
repeat_region 31917..31999
repeat_region /rpt_family="MADE1"
repeat_region 32023..32853
repeat_region /rpt_family="LIM4"
repeat_region 32858..33182
repeat_region /rpt_family="AluJb"
repeat_region 33416..33967
repeat_region /rpt_family="LIM4"
repeat_region complement(33984..34248)
repeat_region /rpt_family="LIP416"
repeat_region 34249..34839
repeat_region /rpt_family="LIP416"
repeat_region 34823..35126
repeat_region /rpt_family="LIM4"
repeat_region 35132..35378
repeat_region /rpt_family="LIM4"
repeat_region 35366..35500
repeat_region /rpt_family="LIM4"
repeat_region complement(35500..35932)
repeat_region /rpt_family="LIP4"
repeat_region 36641..36761
repeat_region /rpt_family="LTR16C"
```

```
alignment_scores:
  Quality: 39.00      Length: 9
  Ratio: 4.333      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 77.778

alignment_block:
US-08-653-294-9 x AC004606/rev ..
Align seg 1/1 to reverse of: AC004606 from: 1 to: 99512
```

```
1 TyrArgLeuAlaIleArgLeuAsnGlu 9
:::|||||:|||||:|||||:|||||:|||||:
5664 TTCAGGCTTCAGTAGATTAATGAG 5638

seq_name: gb_pr3:HSJ1112D6
```

```
seq_documentation_block:
LOCUS HSJ1112D6 135305 bp DNA PRI 23-NOV-1999
DEFINITION Human DNA sequence from clone 1112D6 on chromosome 6q21-22.2,
complete sequence.
ACCESSION AL080317
VERSION AL080317.11 GI:5830430
KEYWORDS HTG; CPG Island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 135305)
```



```

/codon_start=1
/product="annexin V"
/protein_id="BAAL1012.1"
/db_xref="GI:1536796"
/translation="MACLLGAGTVDADPFNDKDEATLHAMKGLGTDDETLKLL
IRSNKROOIALTYKTLFGRLDLDLKLSEKGFETLLVAMVPAHLYDACELRNAI
KGLGTLEWIIEMASRTAAEYKNIRKTYKKEFSDLEKDIVDTSNFERLIVSLVQ
ANDPVGKVGDEVDKALFDAGKNKGTODEETISILSTRGVGHLKRVLDQYMTI
SGVOIEESIQSTGGHFEKLLAVKKSIRSIQCYLAELVLYNSMKGAGTDDQTLIRLV
SRSEIDLFIKRTFRKHYKSLHMIQSDTSGDYNALLLCEIDD"
polyA_signal
1564..1371
BASE COUNT 474 a 312 c 333 g 460 t
ORIGIN

```

```

alignment_scores:
  Quality: 38.00      Length: 10
  Ratio: 3.800      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 60.000

```

```
alignment_block:
US-08-653-294-9 x D64134/rev ..
```

```
Align seg 1/1 to reverse of: D64134 from: 1 to: 1579
```

```

1 TyrArgLeuAlaIleArgLeuAsnGluArg 10
|||||:|||||:|||||:|||||:|||||:
507 TACAGGGTCTCTGTCGCTGACACAGAGA 478

```

```
seq_name: gb_htg4:AC010886
```

```

seq_documentation_block:
LOCUS AC010886 190094 bp DNA HTG 29-OCT-1999
DEFINITION Homo sapiens chromosome unknown clone NH0370K06, WORKING DRAFT
SEQUENCE, in unordered pieces.
ACCESSION AC010886
VERSION AC010886.3 GI:6139236
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 190094)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
2 (bases 1 to 190094)
Waterston,R.
The sequence of Homo sapiens unknown clone NH0370K06
Unpublished
3 (bases 1 to 190094)
Waterston,R.H.
Direct Submission
Submitted (25-SEP-1999) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Oct 29, 1999 this sequence version replaced gi:6015414.
SUBMITTED BY: WUGSC

```

```
COMMENT
```

```

Genome Sequencing Center
Department of Genetics
Washington University
St. Louis MO 63108, USA
http://genome.wustl.edu/gsc
mailto:sapiens@watson.wustl.edu

```

```

NOTICE: This 'working draft' quality sequence may consist of
several contigs from automated sequence assembly concatenated
together. No attempt has been made to order or orient the contigs
relative to one another correctly before concatenating. At each
location in the sequence where contigs have been joined, several
consecutive Ns may have been inserted.

```

The attached annotation was produced using a purely automated procedure.

The location of this clone is unknown.

* NOTE: This is a 'working draft' sequence.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

FEATURES

| Source | Location/Qualifiers |
|---------------|---|
| 1..190094 | /organism="Homo sapiens" |
| | /db_xref="taxon:9606" |
| | /chromosome="unknown" |
| | /clone="NH0370K06" |
| | /clone_lib="unknown" |
| 303..999 | /note="pseudogene similar to PID:g3192702 (AF064539) gpl9 [Bacteriophage N15]" |
| | /note="pseudogene similar to S25023 (PID:g102712) neurofilament-like protein - northern European squid" |
| misc_feature | 2526..2577 /rpt_family="(TG)n" |
| exon | 4528..5741 /rpt_family="Retroviral" |
| misc_feature | 5764..6055 /rpt_family="Retroviral" |
| repeat_region | 6056..6381 /rpt_family="Retroviral" |
| repeat_region | 6382..6416 /rpt_family="(A)n" |
| repeat_region | 6777..6902 /rpt_family="L1" |
| repeat_region | 7313..7436 /rpt_family="MIR" |
| repeat_region | 7506..7738 /rpt_family="MIR" |
| repeat_region | 7741..7806 /rpt_family="(TTA)n" |
| repeat_region | 8013..8465 /rpt_family="L1" |
| repeat_region | 8466..8488 /rpt_family="AT-rich" |
| repeat_region | 9680..9725 /rpt_family="L2" |
| repeat_region | 9778..9805 /rpt_family="AT-rich" |
| repeat_region | 10662..10838 /rpt_family="MIR" |
| repeat_region | 10828..10929 /rpt_family="MIR" |
| repeat_region | 11378..11434 /rpt_family="MIR" |
| repeat_region | 11819..11974 /rpt_family="Retroviral" |
| repeat_region | 12011..12094 /rpt_family="MER53" |
| repeat_region | 12195..12415 /rpt_family="MER2_type" |
| repeat_region | 12466..12556 /rpt_family="Ct-rich" |
| repeat_region | 13205..13252 /rpt_family="(TGAA)n" |
| repeat_region | 13691..13864 /rpt_family="Retroviral" |
| repeat_region | 14039..14378 /rpt_family="MaLR" |
| repeat_region | 14407..14564 /rpt_family="MER1_type" |
| repeat_region | 14648..15582 /rpt_family="L1" |
| repeat_region | 15578..16494 |

```

repeat_region /rpt_family="L1"
16897. 16924
/rpt_family="L2"
16925. 17344
/rpt_family="MaLR"
17345. 17547
/rpt_family="L2"
18834. 18940
/rpt_family="L2"
19127. 19274
/rpt_family="MIR"
19455. 19583
/rpt_family="L2"
19844. 21137
/misc_feature /note="pseudogene similar to PID:g452168 (D21854) H+,K+
-ATPase alpha subunit [Cavia porcellus]"
21140. 21208
/misc_feature /note="pseudogene similar to PID:g452168 (D21854) H+,K+
-ATPase alpha subunit [Cavia porcellus]"
21291. 21377
/rpt_family="L1"
21407. 21515
/rpt_family="CT-rich"
21516. 21741
/rpt_family="L1"
21742. 22029
/rpt_family="L1"
22030. 22059
/rpt_family="L1"
22070. 22315
/rpt_family="Achobo"
22316. 22466
/rpt_family="Alu"
22467. 22564
/rpt_family="Achobo"
22575. 23259
/rpt_family="Achobo"
24227. 24358
/rpt_family="MER1_type?"
24389. 24527
/rpt_family="Retroviral"
24528. 24638
/rpt_family="MER1"
24639. 24842
/rpt_family="Retroviral"
25101. 25578
/rpt_family="MaLR"
25664. 25746
/rpt_family="L2"
26318. 26456
/rpt_family="L2"
28013. 28081
/rpt_family="CA)n"
28083. 28402
/rpt_family="MER2_type"
28542. 29165
/rpt_family="MER1_type"
29166. 30027
/rpt_family="L1"
30028. 31320
/rpt_family="Mariner"
31321. 31348
/rpt_family="L1"
31355. 31383
/rpt_family="(TAAA)n"
31384. 31721
/rpt_family="MER1_type"
31817. 33055
/rpt_family="L1"
33056. 33351
/rpt_family="Alu"
33352. 33799
/rpt_family="L1"

```

```

repeat_region 34630. 34682
/rpt_family="MIR"
34711. 34798
/rpt_family="MIR"
35849. 36154
/rpt_family="(TCTCTG)n"
36173. 36279
/rpt_family="CT-rich"
36543. 36674
/rpt_family="MER103"
37457. 37759
/rpt_family="Alu"
37797. 37816

alignment_scores:
  Quality: 38.00      Length: 10
  Ratio: 3.800       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 70.000

alignment_block:
US-08-653-294-9 x AC010886 ..
Align seg 1/1 to: AC010886 from: 1 to: 190094
1 TTTATGLeuAlaLeuArgLeuAsnGluArg 10
||||:|||||||||||||||||||||:||||:
189288 TATCAGCTAGCTATTAGCTTAATCAGCAA 189317

seq_name: gb_htg3:AC009685

seq_documentation_block:
LOCUS AC009685 210031 bp DNA HTG 29-SEP-1999
DEFINITION Homo sapiens chromosome 15 clone 91_E_13 map 15, *** SEQUENCING IN
PROGRESS ***, 27 unordered pieces.
ACCESSION AC009685
VERSION AC009685.2 GI:5932598
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 210031)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 15, Clone 91_E_13
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 210031)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArrellano,K., Depayre,E., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehotsky,J., Lieu,C., Locke,K., Macdonald,P.,
Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
TITLE
JOURNAL
COMMENT
Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 29, 1999 this sequence version replaced gi:5801723.
All repeats were identified using RepeatMasker: Smit, A.F.A. &
Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

```

* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1
3042: contig of 3042 bp in length
gap of unknown length
3043 5640: contig of 2598 bp in length
gap of unknown length
5641 8607: contig of 2967 bp in length
gap of unknown length
8608 11298: contig of 2691 bp in length
gap of unknown length
11299 14358: contig of 3060 bp in length
gap of unknown length
14359 17979: contig of 3621 bp in length
gap of unknown length
17980 21275: contig of 3296 bp in length
gap of unknown length
21276 26139: contig of 4864 bp in length
gap of unknown length
26140 30631: contig of 4492 bp in length
gap of unknown length
30632 33904: contig of 3273 bp in length
gap of unknown length
33905 38491: contig of 4587 bp in length
gap of unknown length
38492 42794: contig of 4303 bp in length
gap of unknown length
42795 48113: contig of 5319 bp in length
gap of unknown length
48114 52634: contig of 4521 bp in length
gap of unknown length
52635 58829: contig of 6195 bp in length
gap of unknown length
58830 63849: contig of 5020 bp in length
gap of unknown length
63850 69292: contig of 5443 bp in length
gap of unknown length
69293 77023: contig of 7731 bp in length
gap of unknown length
77024 85822: contig of 8799 bp in length
gap of unknown length
85823 97216: contig of 11394 bp in length
gap of unknown length
97217 109917: contig of 12701 bp in length
gap of unknown length
109918 123746: contig of 13829 bp in length
gap of unknown length
123747 139393: contig of 15647 bp in length
gap of unknown length
139394 156571: contig of 17178 bp in length
gap of unknown length
156572 180032: contig of 23461 bp in length
gap of unknown length
180033 203037: contig of 23005 bp in length
gap of unknown length
203038 210031: contig of 6994 bp in length.
Location/Qualifiers
1. .210031
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="91_E_13"
/clone_lib="RPC1-11 Human Male BAC"
BASE COUNT 68335 a 38965 c 38634 g 63180 t 917 others
ORIGIN

```

alignment_scores:
 Quality: 38.00 Length: 9
 Ratio: 4.222 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 77.778

alignment_block:

US-08-653-294-9 x AC009685

Align seg 1/1 to: AC009685 from: 1 to: 210031

```

1 TytArgLeuAlaIleArgLeuAsnGlu 9
||||:||||:||||:||||:||||:
22265 TATAAACTTCAATTAGACTTATGAA 22291

```

seq_name: gb_inl:CEY75B8A

seq_documentation_block:

LOCUS CEY75B8A 298406 bp DNA INV 06-SEP-1999
 DEFINITION Caenorhabditis elegans cosmid Y75B8A, complete sequence.
 ACCESSION AL033514
 VERSION AL033514.1 GI:3873442
 KEYWORDS HTG.
 SOURCE
 ORGANISM
 Caenorhabditis elegans.

Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
 Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
 1 (bases 1 to 298406)

AUTHORS
 Wilson, R., Almscough, R., Anderson, K., Baynes, C., Berks, M.,
 Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J.,
 Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A.,
 Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M.,
 Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laister, N.,
 Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B.,
 O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A.,
 Saunders, D., Showkeen, R., Smaldon, N., Smith, A., Sonhammer, E.,
 Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M.,
 Vaughan, K., Waterston, R., Watson, A., Weinstock, L.,
 Wilkinson-Sproat, J. and Wohlman, P.

2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans

JOURNAL Nature 368 (6466), 32-38 (1994)

MEDLINE 94150718

REFERENCE 2 (bases 1 to 298406)

AUTHORS Barlow, K.

TITLE Direct Submission

JOURNAL Submitted (12-NOV-1998) Louis, MO 63110, USA. E-mail:

jes@sanger.ac.uk or rw@nematoe.wustl.edu

Coding sequences below are predicted from computer analysis, using
 predictions from GeneFinder (P. Green, U. Washington), and other
 available information.

For a graphical representation of this sequence and its analysis
 see:-

http://webace.sanger.ac.uk/cgi-

bin/display?db=wormacesclass=Sequence object=y75B8A

Current sequence finishing criteria for the C. elegans genome
 sequencing consortium are that all bases are either sequenced
 unambiguously on both strands, or on a single strand with both a
 dye primer and dye terminator reaction, from distinct subclones.

Exceptions are indicated by an explicit note.

IMPORTANT: This sequence is NOT necessarily the entire insert of
 the specified clone. It may be shorter because we only sequence
 overlapping sections once, or longer because we arrange for a small
 overlap between neighbouring submissions.

IMPORTANT: This sequence is not the entire insert of clone Y75B8A.
 It may be shorter because we only sequence overlapping sections
 once, or longer because we arrange for a small overlap between
 neighbouring submissions.

The true left end of clone Y75B8 is at 1 in this sequence. The true
 left end of clone Y49E10 is at 298300 in this sequence. The true
 right end of clone Y79H2 is at 184511 in this sequence. The start
 of this sequence (1. .103) overlaps with the end of sequence
 AL110501

The end of this sequence (298300. .298406) overlaps with the start
 of sequence Z98866.
 Location/Qualifiers
 1. .298406
 /organism="Caenorhabditis elegans"
 /db_xref="taxon:6239"

FEATURES
 source

```

/clone="Y75B8A.4"
/notes="predicted using Genefinder; similar to ATPases
associated with various cellular activities (AAA); cDNA
EST EMBL:D32917 comes from this gene; cDNA EST EMBL:D35629
comes from this gene; cDNA EST EMBL:D32378 comes from this
gene; cDNA EST EMBL:D34764 comes from this gene; cDNA EST
y245c10.3 comes from this gene; cDNA EST yk245c10.5 comes
from this gene; cDNA EST yk243a11.3 comes from this gene;
y433c2.3 comes from this gene; cDNA EST
from this gene"
/codon_start=1
/protein_id="CAA22082.1"
/db_xref="GI:3980014"
/translation="MKIEENMELPILVTSGVLLPGASLKIPIRSKLNTQITIEKYLR
SPNCVIVYAKVSDKYVEATYAEVETKATFNSTVHYSLDVLGHRANIDKLSL
PCTIVSKVVDLNEAISNQNAIEKLTGAKIIASNTLDFREIYSLIDKEKYGSLAD
LCVSMQKFLGQWLEFLGANGTDAKVECMKMEKDKANTLKLKPNLSLSPFVD
GKRRKTPNVKNOVEOLEEKLNAIEFSDVSDRVYSELHKLKSMNQCSYNTLMNMLE
LYSSLPWNTSTIDDIELHKAITLIESHAMDDVREVLHKLAVCMNNSVGMILCF
TGGPGIGTSTIAKAIAESMGRKQFQRVSGGIRDESDIRGHRHYVAMPGRITAEKLT
CTNNPVELLDVLYSGNQSPSALELLDPEQNSTFHDHYNIPNIDVSKIMFIA
TANDIRLEPALDRLEIEMSGYICRNVALRAELNSDPGADVLPMVDELPTQISASN
IVAMTEYMEAGVQLERNVGAICRNVALRAELNSDPGADVLPMVDELPTQISASN
IKLILKMKHVKVIEKMRPLPAGVCGSLVTTIGGRVMPLEASKSKGTGIVTGH
LKVYKESILVAKGMLNSERLGLTGLEDQIHVHLPAAGVNDKQSPSAGTGLACALY
SIATNPLRSDAAVTEISLTGHVLPVIGGVKVKVLAQREGLRVVRVLPKASNEEYLKM
DEDIRLEMDVLAETIEDYVIGAMDDKSPVLAKL"
complement(45151..49365)
/gene="Y75B8A.5"
complement(join(45151..45318,45616..45804,47003..47278,
47833..47976,48031..48247,48622..48827,49173..49294,
49341..49365))
/gene="Y75B8A.5"
/notes="similar to Leucine Rich Repeat (2 copies)"
/codon_start=1
/protein_id="CAA22095.1"
/db_xref="GI:3980027"
/translation="MILLIGTVYVOKFNFIWKLILFHLKYSNALRDTCPMGCCOC
EDQVTEGQGVVLPDLPSGYSLEIRNSVRIETKNSFRKMEKLMQIEFENNPN
LGTIKFLAPGLKRLILKFTCPGLTELQNAFSGIQNMGKLIKEFTPIHRIDGH
TRFNQARELTISGEALSRHCANINQLDFLVSGVLYEIPETSTFRFHVHF
KNSQDIPSTSTLSHTLSHLEHSKVPISAPDAFSGLTTOVIELHACQLTISAR
AFANYGELAKILANTIGDLTSSIMSRLKTRIEENTLSCSGCMKMTSVEEMSD
INFCSTASRSIRSFIRAKCLNPSQEISRKTNHLPSISFSSSSSYNYFYLSLVYF
SKTDINNNQRFEDKMLPCRNFTALLPQNLHQISESLIDFTVKFSRKRRSRSEA"
complement(54968..61234)
/gene="Y75B8A.6"
complement(join(54968..55248,56451..56791,57803..57900,
57952..58036,59923..59985,61041..61234))
/gene="Y75B8A.6"
/codon_start=1
/protein_id="CAA22096.2"
/db_xref="GI:5824844"
/translation="NMNMHGMAGMNGPPPPOMLQMHPLPHVAFHAGSSQNSQA
PLQASATPPVPVSSMNRSSQSPQSTREMSHRSTSRRCVGCRCQCKCQCT
YCQDSFQFGGVKQKCLERCLRVENLRORDAPTFAKRVGCNACEDCRQDCQIC
LVCLDRKFENRHFPGAMCAKCRNNAOSIECVSLMSQNPADFVQRAKPRSTELLN
VLSQOQQQSQMSHFQVRYMTGPPPMQMQMQAPQOQIPQPTQPOVOOQHQ
QPNYLSHARLMQAPQVITAPAPILPQCMQMTMPQHQDIDIPYLIPTSFVQAVQ
Q3810..90492
/gene="Y75B8A.7"
join(83810..83880,83950..84130,84998..85432,86585..86742,
87951..88300,89741..90492)
/gene="Y75B8A.7"

alignment_scores:
Quality: 38.00 Length: 10
Ratio: 4.22 Gaps: 0
Percent similarity: 90.000 Percent identity: 70.000
alignment_block:

```

```

US-08-653-294-9 x CEY75B8A/rev ..
Align seg 1/1 to reverse of: CEY75B8A from: 1 to: 298406

1 TyArgLeuAlaIleArgLeuAsnGluArg 10
||||||| :|||:|||||:|||||
203587 TATCGGCTTAAATCGCCTAAATCAACGA 203558

seq_name: gb_htgl1:CEY75B8

seq_documentation_block:
LOCUS CEY75B8 336638 bp DNA HTG 03-DEC-1998
DEFINITION Caenorhabditis elegans chromosome III clone Y75B8, *** SEQUENCING
IN PROGRESS ***, in unordered pieces.
ACCESSION AL022286
VERSION AL022286.1 GI:3702098
KEYWORDS HTG; HTGS-PHASE1.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
REFERENCE 1 (bases 1 to 336638)
AUTHORS Barlow, K.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-1998) Nematode Sequencing Project, Sanger Centre,
Hinxton, Cambridge CB10 1RQ, UK and Department of Genetics,
Washington University, St. Louis, MO 63110, USA. E-mail:
jes@sanger.ac.uk or rwenematode.wustl.edu
COMMENT On Oct 6, 1998 this sequence version replaced gi:3617803.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence. Work on the sequence is in progress
and the release of this data is based on the understanding that the
sequence may change as work continues. The sequence may be
contaminated with foreign sequence from E.coli, yeast, vector,
phage etc. Order of segments is not known; 800 n's separate
segments.
* NOTE: This is a 'working draft' sequence.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

FEATURES
source Location/Qualifiers
1..336638
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/chromosome="III"
/clone="Y75B8"

BASE COUNT 108209 a 59632 c 61379 g 105818 t 1600 others
ORIGIN

alignment_scores:
Quality: 38.00 Length: 10
Ratio: 4.222 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 70.000

alignment_block:
US-08-653-294-9 x CEY75B8/rev ..
Align seg 1/1 to reverse of: CEY75B8 from: 1 to: 336638

1 TyArgLeuAlaIleArgLeuAsnGluArg 10
||||||| :|||:|||||:|||||
203587 TATCGGCTTAAATCGCCTAAATCAACGA 203558

seq_name: gb_sts:G46114

seq_documentation_block:
LOCUS G46114 485 bp DNA STS 23-MAR-1999
DEFINITION Z6550.1 zebrafish AB Danio rerio STS genomic clone Z6550 5',
sequence tagged site.
ACCESSION G46114
VERSION G46114.1 GI:4492405
KEYWORDS STS.

```

```

SOURCE ORGANISM
zebrafish.
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
Cyprinidae; Cyprinidae; Rasbora; Danio.
REFERENCE 1 (bases 1 to 485)
AUTHORS Shimoda, N., Knapik, E.W., Ziniti, J., Sim, C., Yamada, E., Kaplan, S.
and Fishman, M.C.
TITLE A genetic linkage map of the zebrafish with 2000 microsatellite
markers
JOURNAL Unpublished (1998)
COMMENT Contact: Mark C. Fishman
Cardiovascular Research Center
Massachusetts General Hospital
Mail code 1494100A, 149 13th Street, Charlestown, MA 02129, USA
Fax: 6177263806
Email: fishman@mcg.harvard.edu
http://zebrafish.mgh.harvard.edu
Primer A: TGTTTCCTGATTCGCTGTG
Primer B: GCTTACTGTTCTGTCCTCCTG
STS size: 157
PCR Profile:
Presoak: 94 degrees C for 5.0 minutes
Denaturation: 94 degrees C for 1.0 minute
Annealing: 58 degrees C for 1.0 minute
Polymerization: 72 degrees C for 1.5 minute
PCR Cycles: 27
Thermal Cycler: MJ Research PTC-100
Protocol:
Template: 10 ng
each 375 nM
Primer: each 200 uM
dNTPs: each 200 uM
Tag Polymerase: 0.034 units/ul
Total Vol: 10 ul
Buffer:
MgCL2: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3.
FEATURES
source Location/Qualifiers
1..485
/organism="Danio rerio"
/strain="AB"
/db_xref="taxon:7955"
/clone="Z6550"
/clone_lib="zebrafish AB"
/sex="F"
/dev_stage="Adult"
/lab_host="DH5alphaF'10"
/note="Vector: m13mp19 with added BstXI site; V-type:
Phage; Genomic DNA from a single adult zebrafish of AB
strain was digested with AluI, Cae81, HaeIII, NlaVI, or
RsaI. Fragments in the range of 250-500 bp were gel
purified and a BstXI linker was added. The fragments were
cloned into a modified M13mp19 vector and transformed
into E. coli DH5alpha. Microsatellite sequences were
screened with labeled d(CA)15 and d(GT)15 oligonucleotide
probes."
STS 55..211
primer_bind 55..74
primer_bind complement(191..211)
BASE COUNT 107 a 83 c 124 g 130 t 41 others
ORIGIN

alignment_scores:
Quality: 37.00 Length: 8
Ratio: 4.625 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 87.500

alignment_block:

```

US-08-653-294-9 x G46114/rev ..

Align seg 1/1 to reverse of: G46114 from: 1 to: 485

1 TyrArgLeuAlaIleArgLeuAsn 8
|||||
94 TACCGGCTGCCTAGATCAAC 71

seq_name: gb_ro:AB012933

seq_documentation_block:

LOCUS AB012933 2454 bp mRNA 05-FEB-1999
DEFINITION Rattus norvegicus mRNA for acyl-CoA synthetase 5, complete cds.
ACCESSION AB012933
VERSION
KEYWORDS acyl-CoA synthetase 5.
SOURCE Rattus norvegicus cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 2454)

AUTHORS Fujino,T. and Yamamoto,T.

TITLE Direct Submission

JOURNAL Submitted (13-APR-1998) to the DDBJ/EMBL/GenBank databases.
Takahiro Fujino, Tohoku University, Gene Research Center; 1-1
Tsutsumido-1-Amamiya, Sendai, Miyagi 981-0914, Japan
(E-mail:tfujino@biochem.tohoku.ac.jp, Tel: +81-22-717-8875,
Fax: +81-22-263-9295)

2 (sites)

REFERENCE 1 (bases 1 to 2454)

AUTHORS Iijima,H., Suzuki,T., Sasana,H., Sato,H., Kamataki,A.,
Nagura,H., Kang,M.-J., Fujino,T., Suzuki,H. and Yamamoto,T.

TITLE A novel acyl-CoA synthetase, ACS5, expressed in intestinal

JOURNAL epithelial cells and proliferating preadipocytes

MEDLINE J. Biochem. 124 (3), 679-685 (1998)

FEATURES

Source

1. .2454

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

121. .2172

/gene="ACS5"

121. .2172

/gene="ACS5"

/product="acyl-CoA synthetase 5"

/protein_id="BA033581.1"

/db_xref="GI:3721653"

/translation="MLFIFNLFSLPTPALICLTFTGTAIFLMLINRQPVLPLIDL
NQSVIGGARGAFQKNDLILYFSDAKTLYEVFORGLAVSDNGPCLGRKPNQ
YKQSYKQVSDRAEYGLSCLLHKGKPSODQFIGFAQNRPEWISLACTYISMVAV
PLYDLGAEAIYVINRADISVVCIDTPKATMLLENVEKLTPLGKTIVLMDPDDDD
LMKRGKGEIEMLSLHDAENLGENFKAPPNPDELVSVCFTSTGDPGKMLTHQ
NIVSNMAAFLEPIFQPTPEDVTISYLPAMFERLVQVFCGKIGFGQDIIR
LPLDDMKALKPTVFTVPRLLNRVDYQVNEAKTPKRLNLALISKFNEVRGDIR
RNSLMDKLVFSIQSLGKVRMLTGAAPISVPLTFFRAAGCWFEAYGQICBTA
GCSTISPGDMTAGHVGTPVSCNFVLEADVADNVPVSNVEGEICIKGNVFKGLKDP
EXTQVLEKDKHLHGDIGRWLPNGTLIIRKKNIFKLAAGEYIAPKEIENVYSR
PILQVVFHGESLGVVDPESLPFAKIGVKGVSFELCONQKRAILEDLQ
KVGREGGLKSFQVKSIFVHPPEFSIENGLLTPTLTKAKRVELAKEFQIQISLVESIE
E"

BASE COUNT 672 a 571 c 598 g 613 t

ORIGIN

alignment_scores:

Quality: 37.00 Length: 10

Ratio: 3.700 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 60.000

alignment_block:

US-08-653-294-9 x AB012933/rev ..

Align seg 1/1 to reverse of: AB012933 from: 1 to: 2454

1 TyrArgLeuAlaIleArgLeuAsnGluArg 10
|||||
267 TACAGACTGGTTCCTCAAGTCATGAGAG 238

seq_name: gb_p11:YSCCOT1A

seq_documentation_block:

LOCUS YSCCOT1A 2675 bp DNA PLN 25-AUG-1995
DEFINITION Saccharomyces cerevisiae COT1 protein gene, complete cds.
ACCESSION M88252
VERSION
KEYWORDS COT1 protein; transmembrane protein.
SOURCE Saccharomyces cerevisiae DNA.
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycetales;
Saccharomycetaceae; Saccharomycetes.

REFERENCE 1 (bases 1 to 2675)

AUTHORS Conklin,D.S., McMaster,J.A., Culbertson,M.R. and Kung,C.

TITLE COT1, a gene involved in cobalt accumulation in Saccharomyces

JOURNAL Mol. Cell. Biol. 12 (9), 3678-3688 (1992)

MEDLINE 92375034

FEATURES

Source

1. .2675

/organism="Saccharomyces cerevisiae"

/db_xref="taxon:4932"

/chromosome="15"

/tissue_lib="YEp24"

872. .875

/gene="COT1"

/db_xref="SGD:S0005843"

872. .2428

/gene="COT1"

/db_xref="SGD:S0005843"

959. .962

/gene="COT1"

/db_xref="SGD:S0005843"

1003. .2322

/gene="COT1"

/codon_start=1

/db_xref="SGD:S0005843"

/product="COT1 protein"

/protein_id="AAA74884.1"

/db_xref="GI:171263"

/translation="MKLGSKOVKIISLLDITVFEIGTGYLSHSLALIASDFHML
NDIISLVVAVNVNVRNPDSYTYGWRRAEILGALINAVFLIACVSLIALQOR
IIAPPVIEPKFVLVGVAGLISNTVGLFHDNDQEHGHSHGGIFADHEHMP
SSHTTHAHVDGIENTPMDSTNISEIPNAIVDSFNMNTRLLTPENASKTFSYST
SSHTIAGENTYEHKRRSLNMGHGVFLHVGDLGALNIGVMSLFAFFIKTIDYSKYYT
DPLVSLITGIIFSSALPLCKASKILLQATPSLGGDGLKIPGIIAIDFRV
WLVSEFISLHLQDISPEQFDLAKIVRSKLHRYGIHSATIQPEFIEVSTER
AGDSQGHLDNDPLSLRPKTYGTGISTGSLVDVDAANCNTACDLEDH"

misc_feature

/gene="COT1"

/note="hydrophobic domain"

/db_xref="SGD:S0005843"

1129. .1182

/gene="COT1"

/note="hydrophobic domain"

/db_xref="SGD:S0005843"

1234. .1302

/gene="COT1"

/note="hydrophobic domain"

/db_xref="SGD:S0005843"

1342. .1401

/gene="COT1"

/note="hydrophobic domain"

/db_xref="SGD:S0005843"

1420. .1446

/gene="COT1"

/note="potential metal coordination; putative"

/db_xref="SGD:S0005843"


```
misc_feature 1489..1509
              /gene="COT1"
              /note="potential metal coordination; putative"
              /db_xref="SGD:S0005843"
misc_feature 1732..1797
              /gene="COT1"
              /note="hydrophobic domain"
              /db_xref="SGD:S0005843"
misc_feature 1837..1887
              /gene="COT1"
              /note="hydrophobic domain"
              /db_xref="SGD:S0005843"
terminator 2411..2428
            /gene="COT1"
            /db_xref="SGD:S0005843"
BASE COUNT 849 a 521 c 490 g 815 t
ORIGIN
```

```
alignment_scores:
  Quality: 37.00      Length: 10
  Ratio: 3.700       Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 60.000
```

alignment_block:

```
US-08-653-294-9 x YSCCOT1A ..
Align seg 1/1 to: YSCCOT1A from: 1 to: 2675
```

```
1 TyrArgLeuAlaIleArgLeuAsnGluArg 10
||||:||||:||||:||||:||||:
792 TATAAGCTGACGATCCGCATACGAGAAG 821
```

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-9 to: N_Geneseq_36.* out_format : pfs

Date: Feb 8, 2000 1:27 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n_model -DEV=xlp
-Q=/cgnl1/USPTO_spool/US08653294/runat_04022000_160701_15807/app_query.fasta.1
-DB=N_Geneseq_36 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blossum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT
-THREADS=1

Search information block:

Query: US-08-653-294-9
Query length: 10
Database: N_Geneseq_36.*
Database sequences: 311585
Database length: 125096042
Search time (sec): 590.520000

score_list:

| Seq | Strd | Orig | ZScore | Escore | Len | Documentation |
|----------------------|------|-------|--------|---------|-------|----------------------------------|
| N_Geneseq_36:Q20995 | + | 36.00 | 119.69 | 45.71 | 1560 | EHV-4 gC gene. Nucleic acid se |
| N_Geneseq_36:X20531 | + | 36.00 | 111.44 | 131.67 | 3858 | Polynucleotide sequence from t |
| N_Geneseq_36:T59978 | - | 35.00 | 115.78 | 75.47 | 1549 | 3' untranslated region of the |
| N_Geneseq_36:X20218 | - | 35.00 | 115.78 | 75.47 | 1549 | D. melanogaster tipE+ 4kb clon |
| N_Geneseq_36:T59975 | - | 35.00 | 107.24 | 225.60 | 3954 | Full length tipE protein codin |
| N_Geneseq_36:X20215 | - | 35.00 | 107.24 | 225.60 | 3954 | D. melanogaster tipE+ genomic |
| N_Geneseq_36:X51895 | + | 34.00 | 127.46 | 16.88 | 278 | Human secreted protein 5' EST S |
| N_Geneseq_36:X78539 | + | 33.00 | 121.84 | 34.71 | 333 | Staphylococcus aureus contig SE |
| N_Geneseq_36:T33136 | + | 33.00 | 108.86 | 183.39 | 1384 | Broccoli ACC synthase genomic |
| N_Geneseq_36:X20113 | + | 32.00 | 117.78 | 58.40 | 336 | Probe (14) for microbial genes |
| N_Geneseq_36:T37110 | + | 32.00 | 109.37 | 171.78 | 846 | Antibiotic resistance gene blaz |
| N_Geneseq_36:T75169 | + | 32.00 | 106.37 | 252.41 | 1176 | Staphylococcus aureus contig s |
| N_Geneseq_36:T32800 | - | 32.00 | 101.40 | 477.67 | 2030 | Protein disulphide isomerase g |
| N_Geneseq_36:T29774 | - | 32.00 | 101.35 | 480.42 | 2040 | Bacillus thuringiensis ssp. is |
| N_Geneseq_36:N90712 | - | 32.00 | 101.09 | 496.97 | 2100 | cryp gene. Bacillus thuringien |
| N_Geneseq_36:Q14809 | - | 32.00 | 98.46 | 696.11 | 2802 | Btm PG14 72kDa Cry insecticida |
| N_Geneseq_36:Q14810 | - | 32.00 | 98.14 | 724.94 | 2901 | Btm PG14 72kDa Cry insecticida |
| N_Geneseq_36:Q73061 | - | 32.00 | 94.34 | 1.2e+03 | 4403 | SefABCD gene cluster. Elicitin |
| N_Geneseq_36:X13232 | + | 32.00 | 93.67 | 1.3e+03 | 4738 | Enterococcus faecalis genome c |
| N_Geneseq_36:T80415 | - | 32.00 | 83.81 | 4.6e+03 | 13987 | Hybrid smrg/Cyig ORF1. DNA en |
| N_Geneseq_36:T78508 | - | 32.00 | 73.29 | 1.7e+04 | 44377 | Platenolide synthase gene clu |
| N_Geneseq_36:T80412 | + | 32.00 | 73.29 | 1.7e+04 | 44377 | Platenolide synthase gene clu |
| N_Geneseq_36:V64212 | + | 31.00 | 124.67 | 24.14 | 102 | N. gonorrhoeae pilC1 PCR amplifi |
| N_Geneseq_36:V64213 | + | 31.00 | 124.58 | 24.42 | 103 | N. gonorrhoeae pilC1 PCR amplifi |
| N_Geneseq_36:V64214 | + | 31.00 | 124.49 | 24.70 | 104 | N. gonorrhoeae pilC1 PCR amplifi |
| N_Geneseq_36:T77268 | + | 31.00 | 112.36 | 117.10 | 394 | Staphylococcus aureus contig SE |
| N_Geneseq_36:X20596 | + | 31.00 | 106.34 | 253.48 | 763 | Polynucleotide sequence from th |
| N_Geneseq_36:T77280 | + | 31.00 | 105.51 | 281.65 | 835 | Lymantria dispar nuclear polyhe |
| N_Geneseq_36:X202039 | + | 31.00 | 104.21 | 333.12 | 964 | Determination gene for Pasteure |
| N_Geneseq_36:X20619 | + | 31.00 | 103.08 | 384.95 | 1091 | Polynucleotide sequence from b |
| N_Geneseq_36:T74973 | - | 31.00 | 100.82 | 534.31 | 1398 | Staphylococcus aureus contig S |
| N_Geneseq_36:V07921 | - | 31.00 | 96.13 | 937.96 | 2338 | Helicobacter pylori 76 kDa pol |
| N_Geneseq_36:V07916 | - | 31.00 | 95.79 | 980.75 | 2429 | Helicobacter pylori 76 kDa pol |
| N_Geneseq_36:Q85560 | + | 31.00 | 94.54 | 1.2e+03 | 2785 | Ethylene response (ETR) gene c |
| N_Geneseq_36:Q85567 | + | 31.00 | 94.53 | 1.2e+03 | 2787 | Ethylene response (ETR) gene c |
| N_Geneseq_36:Q85561 | + | 31.00 | 94.53 | 1.2e+03 | 2787 | Ethylene response (ETR) gene c |
| N_Geneseq_36:V59032 | + | 31.00 | 94.53 | 1.2e+03 | 2787 | A. thaliana ethylene response |
| N_Geneseq_36:V59034 | + | 31.00 | 94.53 | 1.2e+03 | 2787 | A. thaliana ethylene response |
| N_Geneseq_36:V59035 | + | 31.00 | 94.53 | 1.2e+03 | 2787 | A. thaliana ethylene response |
| N_Geneseq_36:V59036 | + | 31.00 | 94.53 | 1.2e+03 | 2787 | A. thaliana ethylene response |
| N_Geneseq_36:V59037 | + | 31.00 | 94.53 | 1.2e+03 | 2787 | A. thaliana ethylene response |

N_Geneseq_36:Q85556 + 31.00 91.52 1.7e+03 3879 | Ethylene response (ETR) gen
N_Geneseq_36:V59031 + 31.00 91.52 1.7e+03 3879 | A. thaliana ethylene respon
N_Geneseq_36:V74399 - 31.00 91.51 1.7e+03 3886 | Staphylococcus aureus conti
N_Geneseq_36:T49073 - 31.00 90.94 1.8e+03 4136 | Agaricus bisporus hypa and

seq_name: N_Geneseq_36:Q20995

seq_documentation_block:

ID Q20995 standard; DNA: 1560 BP.
AC Q20995;
DE 19-MAY-1992 (first entry)
DT EHV-4 gC gene.
KW Equine herpes virus-4; glycoprotein gC; antigenic; vaccine;
KW alphaherpesvirus; respiratory disease; cellular attachment;
KW pathogenic; ss.
OS Equine herpesvirus-4.
FH Key Location/Qualifiers
FT cds 52..1509
FT /*tag= a
FT /product= EHV-4_gC

PN W09201057-A.

PD 23-JAN-1992.

PR 04-JUL-1991; G01091.

PR 06-JUL-1990; GB-014950.

PA (UNIU) UNIV OF GLASGOW.

PA (EQUI-) EQUINE VIROLOGY RES FOUN.

PI Nicolson L, Onions DE;

DR WPI: 92-056872/07.

DR P-PSDB; R20796.

PT Nucleic acid sequence encoding EHV-4 gH or gC protein - used to
produce a vaccine for protection of horses against EHV-4

PT Infection

PS Claim 1; Page 23; 29pp; English.

CC Equine dermal cells (NBL-6) were infected with EHV-4 strain 1942
viral DNA, purified and a BamHI library constructed in pUC9.

CC Calcium shocked E. coli DH1 cells were transformed with the

CC recombinant plasmids. Additional clones were derived from a

CC restriction digest of pUC9 contg. the BamHI G fragment. The

CC nucleotide sequence of a region of BamHI G fragment spanning the gC

CC gene was determined. By analysis of overlapping sequences (SEQ ID NO

2). Vaccines can be prep'd. using this sequence, and they may be used

CC to protect horses against EHV-4 infection, inducing a higher level

CC of immunity and less side-effects than other live virus vaccines.

CC See also Q20994.

SQ Sequence 1560 BP; 438 A; 409 C; 334 G; 379 T;

alignment_scores:

Quality: 36.00 Length: 10
Ratio: 4.000 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 70.000

alignment_block:

US-08-653-294-9 x Q20995

Align seg 1/1 to: Q20995 from: 1 to: 1560

1 TyrArgLeuAlaileArgLeuAsnGluArg 10

||||||| ||||| ||||| ||||| |||||

409 TACAGACTAGTAATTCACCTAAACGAGCG 438

seq_name: N_Geneseq_36:X20531

seq_documentation_block:

ID X20531 standard; DNA: 3858 BP.

AC X20531;

DT 05-MAY-1999 (first entry)

DE Polynucleotide sequence from the genome of Treponema pallidum.

KW Treponema pallidum infection; syphilis; Borrelia infection; animal;

OS enzyme production; ds.

PN Treponema pallidum.

PD W09859034-A2.

PD 30-DEC-1998.

```

PF 23-JUN-1998; U13041.
PA 24-JUN-1997; US-050667.
PI (HUMA-) HUMAN GENOME SCI INC.
DR Fraser JN;
WPI: 99-081273/07.
PT New isolated Treponema pallidum nucleic acids - used to develop
PT products for the detection, diagnosis, characterisation, prevention
PT and therapy of T. pallidum infections, particularly syphilis
PS Claim 1, Page 368-370; 1150pp; English
CC X20500-21243 represent polynucleotide sequences from the genome of
CC Treponema pallidum. The sequences can be used for detection,
CC diagnosis, characterisation, prevention and therapy for T. pallidum
CC infections, particularly syphilis. They can also be used for detecting
CC diseases related to Borrelia infections in animals, and for the
CC production of biosynthetic products such as enzymes.
SQ Sequence: 3858 BP; 659 A; 777 C; 1306 G; 1116 T;

alignment_scores:
  Quality: 36.00 Length: 9
  Ratio: 4.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 77.778

alignment_block:
US-08-653-294-9 x X20531/rev ..
Align seg 1/1 to reverse of: X20531 from: 1 to: 3858

2 ArgLeuAlaIleArgLeuAsnGluArg 10
|||||:|||||:|||||:|||||:|||||:
2281 CGATCCACCATTTAGCTGAACGAGCGC 2255

seq_name: N_Geneseq_36:T59978

seq_documentation_block:
ID T59978 standard; DNA; 1549 BP.
AC X20531;
DT 12-MAY-1997 (first entry)
DE 3' untranslated region of the tpeE protein coding sequence.
KW Drosophila; tpeE; para protein; voltage-dependent cation channel; stroke;
KW pesticide; insecticide; insect; parasitic infection; human; head trauma;
KW neuroprotection; hypoxia; therapy; ss.
OS Drosophila melanogaster.
PN US593862-A.
PD 14-JAN-1997.
PF 04-OCT-1994; 317880.
PR (UINY ) UNIV NEW YORK STATE RES FOUND.
PI Feng G, Hall LM;
WPI: 97-099467/09.
PT Nucleic acid encoding Drosophila melanogaster tpeE protein - for
PT prodn. of recombinant voltage-dependent cation channel
PS Example 11; Column 37-40; 33pp; English.
CC T59977 and T59978 represent the 5' and 3' untranslated regions of the
CC Drosophila tpeE protein (see W13843) coding sequence (see T59975).
CC Mutations in the tpeE protein, result in a ethyl methane sulphonate-
CC induced recessive mutation phenotype. Homologous files for the mutation
CC paralyse rapidly at 38 degrees, and recover immediately when returned to
CC 23 degrees. Coexpression of the full length tpeE sequence, and a nucleic
CC acid encoding a para protein results in translation products that form a
CC functional voltage-dependent cation channel. The cation channel can be
CC used to screen for pesticides active against insects such as Drosophila
CC melanogaster and pest insects. The cation channel can also be used to
CC screen for drugs for use in the treatment and prevention of parasitic
CC infections in humans and animals, and to screen drugs for their
CC neuroprotective effect against hypoxia, stroke, and head trauma.
SQ Sequence 1549 BP; 597 A; 328 C; 249 G; 375 T;

alignment_scores:
  Quality: 35.00 Length: 9
  Ratio: 3.889 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 66.667

alignment_block:
US-08-653-294-9 x X20518/rev ..
Align seg 1/1 to reverse of: X20518 from: 1 to: 1549

1 TyrArgLeuAlaIleArgLeuAsnGlu 9
|||||:|||||:|||||:|||||:|||||:
889 TTTCGTTTATCATTCATTCGTTTAAATGAA 863

seq_name: N_Geneseq_36:X202018

seq_documentation_block:
ID X202018 standard; DNA; 1549 BP.
AC X202018;
DT 21-APR-1999 (first entry)
DE D. melanogaster tpeE+ 4kb clone 3'-UTR.
KW tpeE+; para protein; modulator; voltage dependent cation channel; VDCC;
KW pesticide; insect control; pharmaceutical agent; neuroprotection;
KW hypoxia; ischaemia; stroke; head trauma; ss.
OS Drosophila melanogaster.
PN Key
FH 3'UTR
FT 1..1549
FT US5871940-A.
FT 16-FEB-1999.
FT 13-JAN-1997; 782396.
PR 04-OCT-1994; US-317880.
PR 13-JAN-1997; US-782396.
PA (UINY ) UNIV NEW YORK STATE RES FOUND.
PI Feng G, Hall LM;
WPI: 99-166632/14.
PT Screening for agents which modulate ion channel function - using
PT host cells transformed with nucleic acids encoding the Drosophila
PT melanogaster tpeE and para proteins
PS Disclosure: Column 37-40; 54pp; English.
CC This sequence represents the 3'-UTR from a Drosophila melanogaster tpeE+
CC protein which is used in a method for screening for agents which modulate
CC ion channel function which uses host cells transformed with nucleic acid
CC encoding Drosophila melanogaster tpeE and para proteins. Co-expression of
CC these genes in the host cell, allows the formation of a functional
CC voltage dependent cation channel (VDCC) in the cell. The agents
CC identified can be used as pesticides for the control of Drosophila
CC melanogaster or other insects. They can also be used to screen
CC pharmaceutical agents for their neuroprotective affect against e.g.
CC hypoxia, ischaemia, stroke and head trauma.
SQ Sequence 1549 BP; 597 A; 328 C; 249 G; 375 T;

alignment_scores:
  Quality: 35.00 Length: 9
  Ratio: 3.889 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 66.667

alignment_block:
US-08-653-294-9 x X202018/rev ..
Align seg 1/1 to reverse of: X202018 from: 1 to: 1549

1 TyrArgLeuAlaIleArgLeuAsnGlu 9
|||||:|||||:|||||:|||||:|||||:
889 TTTCGTTTATCATTCATTCGTTTAAATGAA 863

seq_name: N_Geneseq_36:T59975

seq_documentation_block:
ID T59975 standard; DNA; 3954 BP.
AC T59975;
DT 12-MAY-1997 (first entry)
DE Full length tpeE protein coding sequence.
KW Drosophila; tpeE; para protein; voltage-dependent cation channel; stroke;
KW pesticide; insecticide; insect; parasitic infection; human; head trauma;
KW neuroprotection; hypoxia; therapy; ss.

```

PS Disclosure; Fig 7A-B; 54pp; English.
CC This sequence encodes a Drosophila melanogaster tipE+ protein which is
CC used in a method for screening for agents which modulate ion channel
CC function which uses host cells transformed with nucleic acid encoding
CC Drosophila melanogaster tipE and para proteins. Co-expression of these
CC genes in the host cell, allows the formation of a functional voltage
CC dependent cation channel (VDCC) in the cell. The agents identified
CC can be used as pesticides for the control of Drosophila melanogaster or
CC other insects. They can also be used to screen pharmaceutical agents for
CC their neuroprotective affect against e.g. hypoxia, ischaemia, stroke and
CC head trauma.

SQ Sequence 3954 BP; 1293 A; 960 C; 867 G; 834 T;

alignment_scores:
Quality: 35.00 Length: 9
Ratio: 3.899 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 66.667

alignment_block:
US-08-653-294-9 x X02015/rev ..
Align seg 1/1 to reverse of: X02015 from: 1 to: 3954

seq_name: N_Geneseq_36:X02015

seq_documentation_block:
ID X51895 standard; DNA; 278 BP.
AC X51895;
DT 22-JUN-1999 (first entry)
DE Human secreted protein 5', EST SEQ ID NO: 109.
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition; ds.
OS Homo sapiens.
PN WO9906552-A2.
PD 11-FEB-1999.
PF 31-JUL-1998; IB1236.
PR 01-AUG-1997; US-905223.
PA (GSE) GENSET.
PI Duclert A, Dumas Milne Edwards J, Lacroix B;
DR WPI: 99-153782/13.
DR P-PSDB; Y13095.
PT New isolated brain-derived nucleic acids - used to develop products
PT which may have cytokine, immune, regulatory, haematopoiesis
PT regulating, anti-inflammatory or tumour inhibition activity
PS Claim 1; Page 256; 577pp; English.
CC X51787 to X52019 represent 5' expressed sequence tags (ESTs) for human
CC secreted proteins, and encode the proteins given in Y12987 to Y13219,
CC respectively. The proteins given represent the signal peptide and an
CC N-terminal fragment of a secreted protein. The nucleic acid sequences
CC can be used for producing secreted human gene products. They can also
CC be used to develop products for diagnosis and therapy. The proteins
CC obtained may have cytokine activity, cell proliferation/differentiation
CC activity, haematopoiesis regulating activity, tissue growth regulating
CC activity, reproductive hormone regulating activity, chemotactic/
CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC ligand activity, anti-inflammatory activity, tumour inhibition activity/
CC or other activities. The products can be used in forensic, gene therapy
CC and chromosome mapping procedures. The sequences can also be used for
CC obtaining corresponding promoter sequences. The nucleic acids encoding
CC the signal peptide can be used for directing extracellular secretion of
CC a polypeptide or the insertion of a polypeptide into a membrane, or
CC importing a polypeptide into a cell.
SQ Sequence 278 BP; 95 A; 58 G; 67 T;

OS Drosophila melanogaster.
FH Location/Qualifiers
FT cds
FT /*tag= a
FT /product= tipE
PN
US593862-A.
PD 14-JAN-1997.
PP 04-OCT-1994; 317880.
PR 04-OCT-1994; US-317880.
PA (UYNV) UNIV NEW YORK STATE RES FOUND.
PI Feng G, Hall LM;
DR WPI: 97-099467/09.
DR P-PSDB; W13843.
PT Nucleic acid encoding Drosophila melanogaster tipE protein - for
PT prodn. of recombinant voltage-dependent cation channel
PS Disclosure; Column 29-32; 33pp; English.
CC This sequence represents the full length coding sequence for the
CC Drosophila tipE protein. Mutations in the tipE protein, result in a
CC ethyl methane sulphonate-induced recessive mutation phenotype.
CC Homologous files for the mutation paralyse rapidly at 38 degrees, and
CC recover immediately when returned to 23 degrees. Coexpression of this
CC sequence, and a nucleic acid encoding a para protein results in
CC translation products that form a functional voltage-dependent cation
CC channel. The cation channel can be used to screen for pesticides active
CC against insects such as Drosophila melanogaster and pest insects. The
CC cation channel can also be used to screen for drugs for use in the
CC treatment and prevention of parasitic infections in humans and animals,
CC and to screen drugs for their neuroprotective effect against hypoxia,
CC stroke, and head trauma.
SQ Sequence 3954 BP; 1293 A; 960 C; 867 G; 834 T;

alignment_scores:
Quality: 35.00 Length: 9
Ratio: 3.899 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 66.667

alignment_block:
US-08-653-294-9 x T59975/rev ..
Align seg 1/1 to reverse of: T59975 from: 1 to: 3954

seq_name: N_Geneseq_36:X02015

seq_documentation_block:
ID X02015 standard; DNA; 3954 BP.
AC X02015;
DT 21-APR-1999 (first entry)
DE D. melanogaster tipE+ genomic DNA.
KW tipE+; para protein; modulator; voltage dependent cation channel; VDCC;
KW pesticide; insect control; pharmaceutical agent; neuroprotection;
KW hypoxia; ischaemia; stroke; head trauma; ss.
OS Drosophila melanogaster.
FH Key
FT Location/Qualifiers
FT cds
FT /*tag= a
FT /product= "tipE+"
PN
US5871940-A.
PD 16-FEB-1999.
PP 13-JAN-1997; 782396.
PR 04-OCT-1994; US-317880.
PR 13-JAN-1997; US-782396.
PA (UYNV) UNIV NEW YORK STATE RES FOUND.
PI Feng G, Hall LM;
DR WPI: 99-166632/14.
DR P-PSDB; W92459.
PT Screening for agents which modulate ion channel function - using
PT host cells transformed with nucleic acids encoding the Drosophila
TT melanogaster tipE and para proteins

Drosophila melanogaster.

Key Location/Qualifiers
cds 1046..2404
 /*tag= a

FT /product- tipe

PEN US593862-A.
PD 14-JAN-1997.
PF 04-OCT-1994; 317880.
PR 04-OCT-1994; US-317880.
PA (UNY) UNIV NEW YORK STATE RES FOUND.
PI Feng G, Hall LM;
DR WPI: 97-099467/09.
P-PSDB; WI3843.

PT Nucleic acid encoding Drosophila melanogaster tipE protein - for prodn. of recombinant voltage-dependent cation channel

PS Disclosure: Column 29-32; 33pp; English.

CC This sequence represents the full length coding sequence for the drosophila tipE protein. Mutations in the tipE protein, result in a ethyl methane sulphate-induced recessive mutation phenotype.

CC Homologous files for the mutation paralyze rapidly at 38 degrees, and recover immediately when returned to 23 degrees. Coexpression of this sequence, and a nucleic acid encoding a para protein results in translation products that form a functional voltage-dependent cation channel. The cation channel can be used to screen for pesticides active against insects such as Drosophila melanogaster and pest insects. The cation channel can also be used to screen for drugs for use in the treatment and prevention of parasitic infections in humans and animals, and to screen drugs for their neuroprotective effect against hypoxia, stroke, and head trauma.

CDS Sequence 3954 BP; 1293 A; 960 C; 867 G; 834 T;

alignment_scores:
Quality: 35.00 Length: 9
Ratio: 3.889 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 66.667

alignment_block:
US-08-653-294-9 x T59975/rev ..

Align seg 1/1 to reverse of: T59975 from: 1 to: 3954

seq_name: N_Geneseq_36.X02015

seq_documentation_block:
ID X02015 standard; DNA; 3954 BP.
AC X02015;
DE 21-APR-1999 (first entry)
KW tipE; para protein; modulator; voltage dependent cation channel; VDCC; pesticide; insect control; pharmaceutical agent; neuroprotection; hypoxia; ischemia; stroke; head trauma; ss.
OS Drosophila melanogaster.
FH Key Location/Qualifiers
CDS 1046..2404
 /*tag= a
 /product= "tipE"

FT FT
FT US5871940-A.
PN PD 16-FEB-1999.
PP PF 13-JAN-1997; 782396.
PR PR 13-OCT-1994; US-317880.
PA (UNY) UNIV NEW YORK STATE RES FOUND.
PI Feng G, Hall LM;
DR WPI: 99-166632/14.
P-PSDB; W92459.

PT Screening for agents which modulate ion channel function - using host cells transformed with nucleic acids encoding the Drosophila melanogaster tipE and para proteins

seq_name: N_Geneseq_36.X02015

seq_documentation_block:
ID X02015 standard; DNA; 3954 BP.
AC X02015;
DE 21-APR-1999 (first entry)
KW tipE; para protein; modulator; voltage dependent cation channel; VDCC; peptide; hormone regulating activity; tumour inhibition activity
OS Homo sapiens.
FN WO9906552-A2.
PD 11-FEB-1999.
PE 31-JUL-1998; IB1236.
PR 01-AUG-1997; US-905223.
PA (GIST) GENSET.
PI Duclert A, Dumas Milne Edwards J, Lacroix B;
DR WPI: 99-153782/13.
P-PSDB; YI3095.

PT New isolated brain-derived nucleic acids - used to develop products which may have cytokine, immune, regulatory, haematopoiesis regulating, anti-inflammatory or tumour inhibition activity

PS Claim ; Page 256; 577pp; English.

CC X51787 to X52019 represent 5' expressed sequence tags (ESTs) for human secreted proteins, and encode the proteins given in Y12987 to Y13219, respectively. The proteins given represent the signal peptide and an N-terminal fragment of a secreted protein. The nucleic acid sequences can be used for producing secreted human gene products. They can also be used to develop products for diagnosis and therapy. The proteins obtained may have cytokine activity, cell proliferation/differentiation activity, haematopoiesis regulating activity, tissue growth regulating activity, reproductive hormone regulating activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, receptor/ligand activity, anti-inflammatoary activity, tumour inhibition activity or other activities. The products can be used in forensic, gene therapy and chromosome mapping procedures. The sequences can also be used for obtaining corresponding promoter sequences. The nucleic acids encoding the signal peptide can be used for directing extracellular secretion of a polypeptide or the insertion of a polypeptide into a membrane, or importing a polypeptide into a cell.

CDS Sequence 278 BP; 95 A; 58 G; 67 T;

alignment_scores:
 quality: 34.00 Length: 9
 Ratio: 3.778 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 66.667

alignment_block:

US-08-653-294-9 x X51895 ..

Align seg 1/1 to: X51895 from: 1 to: 278

1 TyrArgLeuAlaIleArgLeuAsnGlu 9
 ||| |||||:||||:||||:
 241 TACGACTAGCTCTCGCATACAG 267

seq_name: N_Geneseq_36:V78539

seq_documentation_block:

ID V78539 standard; DNA; 333 BP.

AC V78539:

DT 16-MAR-1999 (first entry)

DE Staphylococcus aureus Config SEQ ID #4228.

KW Computer; readable medium; vaccine; S.aureus infection; immunodetection;

KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;

KW skin infection; surgical wound infection; scalded skin syndrome;

KW toxic shock syndrome; ds.

OS Staphylococcus aureus.

PN EP-786539-A2.

PD 30-JUL-1997.

PF 07-JAN-1997; 100117.

PR 05-JAN-1996; US-009861.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,

PI Rosen CJ;

DR WPI: 97-374922/35.

PT Polynucleotide(s) and proteins derived from Staphylococcus aureus

PT stored in computer readable medium and used in the production of

PT anti-S.aureus vaccines

PS Claim 1; Page 2879; 3271pp; English.

CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences

CC of the invention. The DNA sequences are recorded on a computer readable

CC medium, preferably selected from a floppy or hard disk, random access

CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using

CC the S.aureus DNA sequences allows putative functions to be assigned so

CC that protein-encoding or regulatory regions of commercial, therapeutic or

CC industrial importance can be obtained. Specifically, sequences which are

CC likely to encode antigens have been identified and these polypeptides can

CC be used in a vaccine composition against S.aureus infection. The

CC polypeptides can also be used in a kit for the immunodetection of

CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,

CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,

CC skin and surgical wound infections, scalded skin syndrome, toxic shock

CC syndrome, etc. Organisms transformed with the DNA sequences can be used

CC for recombinant production of the polypeptides. The new DNA sequences

CC (and their fragments) are useful as primers or probes for isolating

CC homologues of any of the S.aureus DNA sequences contained on the

CC computer readable medium.

SQ Sequence: 333 BP; 90 A; 91 C; 79 G; 70 T;

alignment_scores:
 quality: 33.00 Length: 10
 Ratio: 3.667 Gaps: 0
 Percent Similarity: 90.000 Percent Identity: 60.000

alignment_block:

US-08-653-294-9 x V78539 ..

Align seg 1/1 to: V78539 from: 1 to: 333

1 TyrArgLeuAlaIleArgLeuAsnGluArg 10
 |||||:|||||:||||:
 16 TACGCTTGCTATAGCCATTAAATAAG 45

seq_name: N_Geneseq_36:T33136

seq_documentation_block:

ID T33136 standard; DNA; 1384 BP.

AC T33136:

DT 07-DEC-1996 (first entry)

DE Broccoli ACC synthase genomic DNA clone ACCA1.

KW ACC synthase; 1-aminocyclopropyl-1-carboxylic acid synthase;

KW ethylene; shelf-life; Cucumis melo; melon; transgenic plant;

KW antisense; broccoli; ss.

OS Brassica oleracea.

PH Key Location/Qualifiers

FT primer_bind complement (1..31)

FT /tag= a

FT /note= "primer RMM393"

FT exon 1..37

FT /tag= b

FT /codon_start= 2..4

FT intron 38..134

FT /tag= c

FT exon 135..431

FT /tag= d

FT intron 432..594

FT /tag= e

FT exon 595..1384

FT /tag= f

FT primer_bind 1352..1382

FT /tag= g

FT /note= "primer RMM394"

PN W09621027-A1.

PD 11-JUL-1996.

PF 07-JUN-1995; U07271.

PR 30-DEC-1994; US-366992.

PA (ASGR-) ASGRW SEED CO.

PI Boeshore MI, Carney KJ, Deng RZ, Reynolds JF, Ruttencutter GE;

PI WPI: 96-334002/33.

DR P-PSDB: R98598.

PT DNA encoding 1-amino:cyclo:propyl-1-carboxylic acid synthase of

PT Brassica oleracea - used to regulate ethylene-dependent processes

PT in plants, esp. to improve shelf life

PS Claim 2; Fig1A-B; 50pp; English.

CC Brassica oleracea genomic clone ACCA1 (T33136) codes for

CC 1-aminocyclopropyl-1-carboxylic acid synthase (ACC-synthase)

CC (R98598), an enzyme involved in ethylene biosynthesis. It was

CC obt'd. by subjecting broccoli leaf total genomic DNA to PCR using

CC primers (see also T33137-38) based on the Arabidopsis thaliana ACC

CC synthase gene. The product was cloned into PCR11 to obtain clone

CC ACCA1. Genomic DNA or cDNA (see also T33139) can be inserted, in

CC sense or antisense orientation, into an expression cassette and then

CC transferred to a binary vector suitable for Agrobacterium-mediated

CC plant transformation. The constructs permit control of the level of

CC ACC synthase in a transgenic plant (esp. B. oleracea or Cucumis

CC melo) and hence a control of maturation, ageing and shelf-life.

SQ Sequence 1384 BP; 401 A; 290 C; 310 G; 383 T;

alignment_scores:

Quality: 33.00 Length: 10

Ratio: 3.300 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 50.000

alignment_block:

US-08-653-294-9 x T33136/rev ..

Align seg 1/1 to reverse of: T33136 from: 1 to: 1384

1 TyrArgLeuAlaIleArgLeuAsnGluArg 10
 |||||:|||||:||||:
 601 TATCGAATCTGTATTCAATAATAAACGT 572

seq_name: N_Geneseq_36:V20113

```

seq_documentation_block:
ID V20113 standard; DNA; 336 BP.
AC V20113;
DT 26-JUN-1998 (first entry)
DE Probe (14) for microbial genes induced during host infection.
KW Probe; identification; microbial gene; pathogenic microorganism;
KW host infection; virulence gene; vaccine; antimicrobial agent; ss.
OS Salmonella typhimurium.
PN W09744487-A1.
PD 27-NOV-1997.
PF 16-MAY-1997; U08208.
PR 17-MAY-1996; US-651155.
PA (REGC ) UNIV CALIFORNIA.
PI Conner CP, Heithoff DM, Mahan MJ;
DR WPI; 98-018538/02.
PT Identification of microbial coding sequences - for use in vaccines
PT against virulent pathogenic microorganisms
PS Claim 3; Page 13; 172pp; English.
CC The present sequence, which is derived from a Salmonella
CC typhimurium gene specifically induced during host infection, can be
CC used as a probe for the identification of microbial genes
CC specifically induced in a pathogenic microorganism during host
CC infection, i.e. virulence genes. The products of the identified
CC virulence genes provide targets for the development of vaccines or
CC antimicrobial agents. The vaccines can be used to protect a
CC mammalian host against the pathogenesis of virulent microorganisms.
CC Sequence 336 BP; 74 A; 86 C; 99 G; 77 T;
SQ Sequence 336 BP; 74 A; 86 C; 99 G; 77 T;

alignment_scores:
Quality: 32.00 Length: 10
Ratio: 3.556 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 70.000

alignment_block:
US-08-653-294-9 x V20113 ..
Align seg 1/1 to: V20113 from: 1 to: 336

1 TyrArgLeuAlaIleArgLeuAsnGlu 10
:::|||||
50 TTCGTTTGGCAATACGACTAAAGATGCG 79

seq_name: N_Geneseq_36:V37110

seq_documentation_block:
ID V37110 standard; DNA; 846 BP.
AC V37110;
DT 04-SEP-1998 (first entry)
DE Antibiotic resistance gene blaZ for beta-lactams.
KW Detection; bacterial antibiotic resistance gene; bacteria;
KW fungal species; identification; beta-lactam; ds.
OS Enterococcus sp.
PN W09820157-A2.
PD 14-MAY-1998.
PF 04-NOV-1996; CA0829.
PR 04-NOV-1996; US-743637.
PA (ID11-) IDI INFECHIO DIAGNOSTIC INC.
PI Bergeron MG, Ouellette M, Picard FJ, Roy PH;
DR WPI; 98-286967/25.
PT Use of oligo:nucleotide primers and probes - for detection.
PT identification and quantification of bacteria, fungi and bacterial
PT antibiotic resistance gene(s)
PS Claim 19; Page 108; 167pp; English.
CC The present sequence represents an antibiotic resistance gene blaZ
CC (resistant to beta-lactams), and was used to design PCR primers
CC V37050-51. The specification describes the use of probes and/or
CC amplification primers which are specific, ubiquitous and sensitive for
CC determining the presence and amount of nucleic acids from a bacterial
CC antibiotic resistance gene and specific bacterial and fungal species in
CC any sample suspected of containing the bacterial or fungal nucleic acids,
CC where each of the nucleic acid or variant or part comprises a selected
CC target region hybridisable with the probes or primers. The method of

```

use comprises contacting the sample with the probes or primers and
 detecting the presence of hybridised probes or amplified products as an
 indication of the presence of the specific bacterial or fungal species
 and bacterial antibiotic resistance genes. The methods and products can
 be used to detect and identify the bacterial and fungal species and
 genera and determine the bacterial resistance to antibiotics.
 Sequence 846 BP; 356 A; 99 C; 135 G; 256 T;

alignment_scores:
 Quality: 32.00 Length: 9
 Ratio: 3.556 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 55.556

alignment_block:
 US-08-653-294-9 x V37110/rev ..

Align seg 1/1 to reverse of: V37110 from: 1 to: 846

1 TyrArgLeuAlaIleArgLeuAsnGlu 9
 :::|||||
 744 TTCAGATGGCCCTAGATTAACAAA 718

seq_name: N_Geneseq_36:V75169

seq_documentation_block:
 ID V75169 standard; DNA; 1176 BP.
 AC V75169;
 DT 16-MAR-1999 (first entry)
 DE Staphylococcus aureus contig SEQ ID #858.
 KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.
 OS Staphylococcus aureus.
 PN EP-786519-A2.
 PD 30-JUL-1997.
 PF 07-JAN-1997; 100117.
 PR 05-JAN-1996; US-009861.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
 PI Rosen CA;
 DR WPI; 97-374922/35.
 PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -
 PT stored on computer readable medium and used in the production of
 PT anti-S.aureus vaccines
 PS Claim 1; Page 1712; 3271pp; English.
 CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S.aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against S.aureus infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S.aureus DNA sequences contained on the
 CC computer readable medium.
 SQ Sequence 1176 BP; 432 A; 167 C; 221 G; 354 T;

alignment_scores:
 Quality: 32.00 Length: 10
 Ratio: 4.571 Gaps: 0
 Percent Similarity: 70.000 Percent Identity: 70.000

```

alignment_block:
US-08-653-294-9 x V75169
..
Align seg 1/1 to: V75169 from: 1 to: 1176

1 TyrArgLeuAlaIleArgLeuAsnGluArg 10
||||| ||||| ||||| ||||| |||||
114 TATTATTGTAATACGACGAATGAGCGA 143

seq_name: N_Geneseq_36:V22800
seq_documentation_block:
ID V22800 standard; DNA; 2030 BP.
AC V22800;
DE 24-JUL-1998 (first entry)
KW Protein disulphide isomerase gene derived from Candida boidinii.
KW Protein disulphide isomerase; PDI; methylotrophic yeast;
KW Candida boidinii strain S2; endoplasmic reticulum; stable conformation;
KW recombination; disulphide bond; secretory protein; gene therapy;
KW endoplasmic reticulum retention signal sequence; genetic engineering; ds.
OS Candida boidinii.
FH Key Location/Qualifiers
FT CDS 367..1962
FT /tag= a
FT /product= PDI
PN BP-828064-A2.
PD 11-MAR-1998.
PF 04-SEP-1997; 306871.
PR 04-SEP-1996; JP-234287.
PA (SUNR ) SUNTORY LTD.
PI Kato N, Sakai Y, Shibano Y;
DR WPI; 98-161102/15.
DR P-PSDB; W56311.
PT Methylotrophic yeast protein di-sulphide isomerase - and
PT corresponding gene useful for increasing yields of secreted
PT heterologous proteins in Candida boidinii
PS Claim 4; Pages 9-12; 30pp; English.
CC The present sequence encodes a protein disulphide isomerase (PDI) derived
CC from a methylotrophic yeast, Candida boidinii strain S2. PDI is a major
CC protein present in the lumen of the endoplasmic reticulum. PDI is
CC believed to be an enzyme which catalyses formation of stable
CC conformation by recombining disulphide bonds of secretory proteins. As
CC the protein must stay in the endoplasmic reticulum, it has a sequence
CC known as the endoplasmic reticulum retention signal sequence at the
CC carboxy terminus (W56317). The PDI protein is useful in gene therapy and
CC genetic engineering. The PDI gene may be coexpressed with a gene of
CC interest to ensure the production of a correctly folded biologically
CC active protein.
SQ Sequence 2030 BP; 678 A; 351 C; 304 G; 697 T;

alignment_scores:
Quality: 32.00 Length: 8
Ratio: 4.571 Gaps: 0
Percent Similarity: 87.500 Percent Identity: 87.500

alignment_block:
US-08-653-294-9 x V22800/rev
..
Align seg 1/1 to reverse of: V22800 from: 1 to: 2030

2 ArgLeuAlaIleArgLeuAsnGlu 9
||||| ||||| ||||| ||||| |||||
1111 CGTTAGATTCGATTAATGAG 1088

seq_name: N_Geneseq_36:T29774
seq_documentation_block:
ID T29774 standard; DNA; 2040 BP.
AC T29774;
DE 19-NOV-1996 (first entry)
DE Bacillus thuringiensis ssp. israelensis CryIVD protein DNA.
KW CryIVD; toxic protein; crystal toxin; expression construct;

transformed cyanobacteria; phycocyanin beta; cpcB; promoter;
insecticide; dipteran larvae; mosquito; blackfly; ss.
Bacillus thuringiensis.
FH Key Location/Qualifiers
FT CDS 41..1972
FT /tag= a
FT US5518897-A.
PN 21-MAY-1996.
PD 21-MAY-1996.
PF 04-MAY-1992; 877876.
PR 04-MAY-1992; US-877876.
PS 28-JAN-1994; US-188581.
PA (UYME-) UNIV MEMPHIS STATE.
PI Murphy RC; Stevens SE;
DR WPI; 96-259063/26.
DR P-PSDB; R97735.
PT New DNA construct for expressing cryIV D protein in cyanobacteria -
PT under control of a phycocyanin beta promoter, useful for control of
PT dipteran larvae in water
PS Example 1; Columns 9-14; 20pp; English.
CC The present sequence encodes the B. thuringiensis ssp. israelensis
CC CryIVD toxic protein, which was used in the prepn. of a claimed DNA
CC construct for the expression of CryIVD in cyanobacteria, comprising
CC the present sequence under the control of phycocyanin beta (cpcB)
CC promoter. Cyanobacteria (which may be adapted for growth in fresh
CC or brackish water) transformed with the construct can be used
CC as insecticides for controlling dipteran larvae, esp. those of
CC mosquitoes and blackflies, that live in water. The percentage of
CC Culex pipiens (mosquito) larvae surviving after 4 days with
CC cyanobacteria transformed with the claimed DNA construct as their
CC only food source was 51 %, compared to 94 % for those fed with
CC cyanobacteria transformed with an empty plasmid. In the
CC cyanobacteria, CryIVD is efficiently expressed under the control of
CC the strong cpcB promoter, even though the cryIVD gene contains 19
CC AUA which are generally poorly translated (if at all) in
CC cyanobacteria.
SQ Sequence 2040 BP; 716 A; 315 C; 373 G; 636 T;

alignment_scores:
Quality: 32.00 Length: 9
Ratio: 4.000 Gaps: 0
Percent Similarity: 88.889 Percent Identity: 66.667

alignment_block:
US-08-653-294-9 x T29774/rev
..
Align seg 1/1 to reverse of: T29774 from: 1 to: 2040

1 TyrArgLeuAlaIleArgLeuAsnGlu 9
||||| ||||| ||||| ||||| |||||
542 TTCGGTCTTTTATAAAGTTAATGAA 516

seq_name: N_Geneseq_36:N90712
seq_documentation_block:
ID N90712 standard; DNA; 2100 BP.
AC N90712;
DE 09-JAN-1990 (first entry)
DE cryD gene.
KW cryD protein; Bacillus thuringiensis; biopesticide.
OS Bacillus thuringiensis var. israelensis.
PN W08907605-A.
PD 24-AUG-1989.
PF 17-FEB-1989; U006663.
PR 19-FEB-1988; US-158176.
PA (ECOG) Ecogen Inc.
PI Donovan WP;
DR WPI; 89-263682/36.
DR P-PSDB; P91462.
PT Bacillus thuringiensis var israelensis cry D toxin gene and proteins
PT - used for producing insecticide compns. active against Dipteran species.
PS Claim 1; fig 2; 58pp; English.
CC cryD gene is inserted into plasmid and used to transform a microorganism.

```


CC The 67kD protein encoded by the gene has insecticidal activity against
CC dipteran larvae.
SQ Sequence 2100 BP; 746 A; 316 C; 378 G; 660 T;

alignment_scores:
Quality: 32.00 Length: 9
Ratio: 4.000 Gaps: 0
Percent Similarity: 88.889 Percent Identity: 66.667

alignment_block:

US-08-653-294-9 x N90712/rev ..

Align seg 1/1 to reverse of: N90712 from: 1 to: 2100

1 TyrArgLeuAlaIleArgLeuAsnGlu 9
:::||||| ||:::|||||||
542 TTCCTCTTTTAAATAAAGTTAAATGAA 516

THIS PAGE BLANK (USPTO)

| ScoreRank | Strd | Orig | ZScore | EScore | Len | Documentation |
|--------------------|------|-------|--------|--------|------|-----------------------------------|
| Sequence | | | | | | |
| gb_gss9:AW146935 | - | 41.00 | 153.04 | 9.28 | 455 | AQ146935 HS_2248_A2_B10_MR_C1T |
| gb_gss9:AW156527 | - | 38.00 | 143.24 | 32.63 | 342 | AQ156527 se28c06.y1 Gm-c1015 G1 |
| gb_est40:AW147227 | - | 37.00 | 142.24 | 37.11 | 235 | AW147227 da16c03.x1 normalized |
| gb_est40:AW147713 | - | 37.00 | 141.87 | 38.91 | 246 | AW147713 da16c03.y1 normalized |
| gb_gss3:B33379 | - | 37.00 | 137.39 | 69.09 | 428 | B33379 HS-1017-A2-F12-NF.abi.C1T |
| gb_gss3:B33467 | - | 37.00 | 133.82 | 109.28 | 666 | B33467 RPC111-512-AVB.RTC11.H |
| gb_gss14:AQ543638 | - | 37.00 | 133.33 | 116.27 | 707 | AQ543638 RPC1-11-346H13.TV.RC1T |
| gb_gss4:AO705201 | - | 36.00 | 132.90 | 122.98 | 453 | AO705201 HS_5522_B2_B06_T7A.RPC |
| gb_est11:AA070757 | - | 36.00 | 131.25 | 151.81 | 555 | AA070757 va67e06.r1 Soares.mous |
| gb_gss4:AO735952 | - | 36.00 | 130.93 | 158.34 | 578 | AO735952 HS_2261_A1_F09_T7C.C1T |
| gb_gss15:AO591361 | - | 36.00 | 130.84 | 160.04 | 584 | AO591361 HS_5410_B2_E01_T7A.RPC |
| gb_est27:AI463881 | - | 36.00 | 130.83 | 160.33 | 585 | AI463881 va67e06.y1 Soares.mous |
| gb_est32:AI728666 | + | 36.00 | 130.28 | 171.99 | 626 | AI728666 BNLC111358 Six-day.CC |
| gb_gss3:BI21136 | - | 36.00 | 124.86 | 344.71 | 1224 | BI21136 F5L16-Sp6 IGF Arabidops |
| gb_est40:AW242923 | - | 35.00 | 133.80 | 109.58 | 246 | AW242923 AV242923.3 riken.full.le |
| gb_est6:W04025 | - | 35.00 | 132.11 | 136.02 | 303 | W04025 T1958 WMTA4 bloodstream |
| gb_est44:AO081183 | - | 35.00 | 132.06 | 136.95 | 305 | AO081183 AO081183 Oncorhynchus |
| gb_est3:R38349 | - | 35.00 | 131.93 | 139.28 | 310 | R38349 vc94a07.s1 Soares.infant |
| gb_gss6:AQ086481 | - | 35.00 | 128.93 | 204.51 | 449 | AQ086481 HS_5543_B2_A12_Sp6E.RP |
| gb_gss10:AQ178284 | - | 35.00 | 128.06 | 228.64 | 500 | AQ178284 HS_2219_B1_A05_T7C.C1T |
| gb_gss4:AO685122 | - | 35.00 | 127.84 | 235.28 | 514 | AO685122 HS_2160_A1_F05_T7C.C1T |
| gb_gss15:AO636347 | - | 35.00 | 127.36 | 250.01 | 545 | AO636347 RPC1-11-478N2.TV.RC1T |
| gb_gss10:AO137379 | + | 35.00 | 125.72 | 308.75 | 668 | AO137379 nbexb0009118r.CUGI.Rice |
| gb_gss6:AQ056431 | - | 35.00 | 124.35 | 367.88 | 791 | AQ056431 nbexb0003115r.CUGI.Rice |
| gb_est22:AI071976 | - | 34.00 | 135.04 | 93.41 | 128 | AI071976 nl-R-C2-nl.g-06-0-UI.s |
| gb_est22:AI018355 | - | 34.00 | 133.65 | 111.63 | 152 | AI018355 ov41a09.s1 Soares_test |
| gb_est8:AA027211 | + | 34.00 | 130.89 | 159.16 | 214 | AA027211 zeg9410.r1 Soares.feta |
| gb_est24:AI243290 | - | 34.00 | 130.85 | 159.94 | 215 | AI243290 q335h07.x1 Soares_NFL |
| gb_est14:AA401730 | - | 34.00 | 129.69 | 185.46 | 248 | AA401730 zt65c10.s1 Soares_test |
| gb_est22:AI004559 | - | 34.00 | 128.10 | 227.49 | 302 | AI004559 zt65h11.s1 Soares_test |
| gb_est15:AA523625 | - | 34.00 | 127.61 | 242.35 | 321 | AA523625 ni71b04.s1 NCI.CGAP.Pr |
| gb_est27:AI435985 | - | 34.00 | 127.14 | 257.24 | 340 | AI435985 zt75d02.s1 Soares_NHMM |
| gb_est15:NA089711 | - | 34.00 | 127.05 | 260.38 | 344 | NA089711 aa50f06.s1 NCI.CGAP.GC |
| gb_est17:AA939302 | - | 34.00 | 126.84 | 267.44 | 353 | AA939302 nr19d09.s1 NCI.CGAP.Pr |
| gb_est2:FI0887 | - | 34.00 | 126.77 | 269.80 | 356 | FI0887 HSC3C072 normalized.Inp |
| gb_est21:AA967396 | - | 34.00 | 126.46 | 280.81 | 370 | AA967396 ua95c07.r1 Soares_2NBd |
| gb_est36:AA189368 | - | 34.00 | 126.31 | 286.32 | 377 | AA189368 w085c07.x1 NCI.CGAP.Ub |
| gb_gss10:AQ222996 | - | 34.00 | 126.14 | 292.62 | 385 | AQ222996 HS_2017_A2_E09_T7.C1T |
| gb_est20:NA0844239 | - | 34.00 | 125.93 | 300.51 | 395 | NA0844239 ai39d07.s1 Soares.para |
| gb_gss1:CN500236 | - | 34.00 | 125.85 | 303.66 | 399 | AL0084468 Arabidopsals.thaliana.9 |
| gb_est24:AI176387 | + | 34.00 | 125.71 | 309.19 | 406 | AI176387 EST219371 Normalized.Inp |
| gb_est26:AI1833239 | + | 34.00 | 125.69 | 309.98 | 407 | AI1833239 ztcf507.x1 Soares_NHMM |

```

SOURCE      soybean.
ORGANISM    Glycine max

REFERENCE   1 (bases 1 to 342)
AUTHORS     Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,
            Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
            Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
            Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
            Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
            Public Soybean EST Project
            Unpublished (1999)
            On Dec 20, 1995 this sequence version replaced gi:1134478.
            Contact: Shoemaker R/Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@wustl.wustl.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134 For further information
            call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
            427-3324 or contact: clones@genomesystems.com or
            info@genomesystems.com web site: www.genomesystems.com
            Seq primer: -40RP from Gibco
            High quality sequence stop: 341.

FEATURES   1..342
            Location/Qualifiers
            source
            1..342
            /organism="Glycine max"
            /db_xref="taxon:3847"
            /clone="GENOME SYSTEMS CLONE ID: Gm-cl015-2531"
            /clone_lib="Gm-cl015"
            /tissue_type="Mature flowers, field grown plants"
            /lab_host="X110-Gold"
            /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
            XhoI; This cDNA library was constructed from mRNA isolated
            from mature flowers of field grown plants. The cDNA
            library was prepared using the Stratagene pBluescript II
            XR cDNA library construction kit. Complementary DNA was
            synthesized from mRNA using a primer consisting of a poly
            (dT) sequence with a XhoI restriction site. EcoRI adapters
            were ligated to the blunt-ended cDNA fragments followed by
            XhoI digestion. The cDNA fragments were directionally
            cloned into the EcoRI-XhoI restriction site of the
            pBluescript vector. The ligated cDNA fragments were
            transformed into X110-Gold host cells. This library was
            constructed by Dr. Randy Shoemaker and Dr. John
            Erpelting."
BASE COUNT  139 a 62 c 49 g 92 t
ORIGIN
alignment_scores
  Quality: 38.00 Length: 9
  Ratio: 4.222 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 77.778

alignment_block
US-08-653-294-9 x AW156527/rev ..
Align seg 1/1 to reverse of: AW156527 from: 1 to: 342

1 TyrArgLeuAlaIleArgLeuAsnGlu 9
81 TATCGTTTGGCATGAATGATGAA 55

seq_name: gb_est40:AW147227
seq_documentation_block:

```

```

LOCUS      AW147227 235 bp mRNA EST 30-NOV-1999
DEFINITION dal6c03.x1 normalized xenopus laevis gastrula xenopus laevis cDNA
            clone XENOPUS_SOURCE_ID: xinga001n05 3', mRNA sequence.
ACCESSION  AW147227
VERSION    AW147227.1 GI:6195123
KEYWORDS   EST.
SOURCE     African clawed frog.
ORGANISM   Xenopus laevis
REFERENCE  1 (bases 1 to 235)
AUTHORS    Johnson,S.L., Blumberg,B., Song,J., Clifton,S., Hillier,L.,
            Pape,D., Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y.,
            Person,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
            Waterston,R. and Wilson,R.
            WashU Xenopus EST project, 1999
            Unpublished (1999)
            On Dec 20, 1995 this sequence version replaced gi:1135577.
            Other_ESTs: dal6c03.yl
            Contact: Stephen L. Johnson/WashU Xenopus EST project, 1999
            WashU Xenopus EST project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@wustl.wustl.edu
            Library constructed by Bruce Blumberg
            Library normalized by Jihwan Song
            DNA sequencing by: Washington University Genome Sequencing Center
            Clone distribution: Xenopus clone distribution information for
            this library can be found through Research Genetics, visit their
            web page at: http://www.resgen.com/
            Seq primer: -40UP from Gibco.

FEATURES   1..235
            Location/Qualifiers
            source
            1..235
            /organism="Xenopus laevis"
            /db_xref="taxon:8355"
            /clone="XENOPUS_SOURCE_ID: xinga001n05"
            /clone_lib="normalized xenopus laevis gastrula"
            /tissue_type="gastrula (stages 10.5, 11.5 mixed)"
            /lab_host="Top-10 F"
            /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
            XhoI; cDNA was prepared from 2ug of poly A+ RNA (equal
            parts from stage 10.5 and stage 11.5 gastrulae).
            EcoRI-XhoI cut cDNA was then ligated into UniZap-XR
            (Stratagene) with EcoRI at the 5' end and XhoI at the 3'
            end. SS-library phagemids were prepared by mass excision
            from the original library and normalized by hybridization
            to biotinylated driver (prepared from the same library by
            PCR) to Cot-omega of 11. After removal of hybrids and
            excess driver by streptavidin sepharose chromatography,
            the ss-phagemids were made double stranded and
            electroporated into Top-10 F'. Original library
            construction by Bruce Blumberg (Cho et al. 1991 Cell 67,
            1111-1120). Normalized by Jihwan Song (Song, Cho and
            Blumberg, unpublished)."
BASE COUNT  70 a 52 c 45 g 68 t
ORIGIN
alignment_scores
  Quality: 37.00 Length: 9
  Ratio: 4.111 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 77.778

alignment_block
US-08-653-294-9 x AW147227/rev ..
Align seg 1/1 to reverse of: AW147227 from: 1 to: 235

1 TyrArgLeuAlaIleArgLeuAsnGlu 9
|||||

```

205 TACAGATTGCCATACGTGTGAACAAG 179

seq_name: gb_est40:AW147713

seq_documentation_block: 246 bp mRNA EST 30-NOV-1999
 LOCUS AW147713
 DEFINITION da16c03.y1 normalized Xenopus laevis gastrula Xenopus laevis cDNA
 clone XENOPUS_SOURCE_ID:xlnga001n05 5', mRNA sequence.

ACCESSION AW147713

VERSION AW147713.1 GI:6195609

KEYWORDS EST.

SOURCE African clawed frog.

ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
 Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae;
 Xenopus.

REFERENCE 1 (bases 1 to 246)

AUTHORS Johnson,S.L., Blumberg,B., Song,J., Clifton,S., Hillier,L.,
 Pape,D., Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y.,
 Person,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
 Waterston,R. and Wilson,R. 1999

TITLE WashU Xenopus EST project, 1999

JOURNAL Unpublished (1999)

COMMENT Other ESTs: da16c03.x1 this sequence version replaced gi:2059622.

Contact: Stephen L. Johnson/WashU Xenopus EST project, 1999

WashU Xenopus EST project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library constructed by Bruce Blumberg

Library normalized by Jihwan Song

DNA Sequencing by: Washington University Genome Sequencing Center

Clone Distribution: Xenopus clone distribution information for

this library can be found through Research Genetics, visit their

web page at: <http://www.resgen.com/>

Seq primer: -40RP from Gibco.

FEATURES Location/Qualifiers

1..246

/organism="Xenopus laevis"

/db_xref="taxon:8355"

/clone="XENOPUS_SOURCE_ID:xlnga001n05"

/clone_lib="normalized Xenopus laevis gastrula"

/tissue_type="gastrula (stages 10.5, 11.5 mixed)"

/lab_host="Top-10 F"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:

XhoI; cDNA was prepared from 2ug of poly A+ RNA (equal

parts from stage 10.5 and stage 11.5 gastrulae).

EcoRI-XhoI cut cDNA was then ligated into Unizap-XR

(Stratagene) with EcoRI at the 5' end and XhoI at the 3'

end. SS-library phagemids were prepared by mass excision

from the original library and normalized by hybridization

to biotinylated driver (prepared from the same library by

PCR) to Cor-omega of 11. After removal of hybrids and

excess driver by streptavidin sepharose chromatography,

the ss-phagemids were made double stranded and

electroporated into Top-10 F'. Original library

constructed by Bruce Blumberg (Cho et al. 1991 Cell 67,

1111-1120). Normalized by Jihwan Song (Song, Cho and

Blumberg, unpublished)."

BASE COUNT 63 a 52 c 59 g 72 t

ORIGIN

alignment_scores: Quality: 37.00 Length: 9

Ratio: 4.111 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 77.778

alignment_block:

US-08-653-294-9 x AW147713

Align seg 1/1 to: AW147713 from: 1 to: 246

1 TyrArgLeuAlaIleArgLeuAsnGlu 9

|||||

63 TACAGATTGCCATACGTGTGAACAAG 89

seq_name: gb_gss3:B33379

seq_documentation_block: 428 bp DNA GSS 17-OCT-1997

LOCUS B33379

DEFINITION HS-1017-A2-F12-MF.abi CIT Human Genomic Sperm Library C Homo

sapiens genomic clone Plate-CT 792 Col=24 Row=K, genomic survey

sequence.

ACCESSION B33379

VERSION B33379.1 GI:2532748

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 428)

AUTHORS Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,

Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.

Construction of a Characterized Clone Resource for Genomic

Sequencing: Generation and Preliminary Analysis of 20,000 Sequence

Tagged Connectors

Unpublished (1997)

COMMENT Contact: Mahairas GG, Zackrone KD, Hood L

University of Washington

Seattle, WA 98195, USA

Tel: (206) 616-8744

Fax: (206) 685-7301

Email: kzackrone@u.washington.edu

Sequence Tagged Connector

Plate: CT 792 row: K column: 24

Class: BAC ends

High quality sequence stop: 428.

FEATURES Location/Qualifiers

1..428

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="Plate-CT 792 Col=24 Row=K"

/clone_lib="CIT Human Genomic Sperm Library C"

/sex="M"

/note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in

E-Coli DH10B"

BASE COUNT 102 a 116 c 100 g 110 t

ORIGIN

alignment_scores: Quality: 37.00 Length: 9

Ratio: 4.111 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 77.778

alignment_block:

US-08-653-294-9 x B33379/rev ..

Align seg 1/1 to reverse of: B33379 from: 1 to: 428

2 ArgLeuAlaIleArgLeuAsnGluArg 10

|||||

405 CGCCTGCTGTACGCTGATTAAGCGA 379

seq_name: gb_gss3:B63467

seq_documentation_block: 666 bp DNA GSS 08-APR-1999

LOCUS B63467

DEFINITION RPC111-5H2.TVB RPC1-11 Homo sapiens genomic clone RPC1-11-5H2,

genomic survey sequence.

ACCESSION B63467

VERSION B63467.1 GI:2637457

KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 666)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search.html
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1. .666
/organism="Homo sapiens"
/db_xref="GDB:7501705"
/db_xref="taxon:9606"
/clone="RPCI-11-5H2"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"
BASE COUNT 256 a 120 c 126 g 164 t
ORIGIN
alignment_scores:
Quality: 37.00 Length: 8
Ratio: 4.625 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-653-294-9 x B63467/rev ..
Align seg 1/1 to reverse of: B63467 from: 1 to: 666
3 LeuAlaIleArgLeuAsnGluArg 10
|||||
543 TTAGCTATCAGGCTTAATGAGAGG 520
seq_name: gb_gss14:AQ543638
seq_documentation_block: 707 bp DNA GSS 19-MAY-1999
LOCUS AQ543638
DEFINITION RPCI-11-346H13-TV RPCI-11 Homo sapiens genomic clone
RPCI-11-346H13, genomic survey sequence.
ACCESSION AQ543638
VERSION AQ543638.1 GI:4873922
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 707)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building

JOURNAL
COMMENT

Unpublished (1997)
Other_GSSs: RPCI-11-346H13.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search.html.
Seq primer: T7
Class: BAC ends.

FEATURES
Source

Location/Qualifiers
1. .707
/organism="Homo sapiens"
/db_xref="GDB:7632660"
/db_xref="taxon:9606"
/clone="RPCI-11-346H13"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"
BASE COUNT 222 a 125 c 123 g 237 t
ORIGIN

alignment_scores:
Quality: 37.00 Length: 9
Ratio: 4.625 Gaps: 0
Percent Similarity: 88.889 Percent Identity: 77.778

alignment_block:

US-08-653-294-9 x AQ543638/rev ..

Align seg 1/1 to reverse of: AQ543638 from: 1 to: 707

1 TyrArgLeuAlaIleArgLeuAsnGlu 9
|||||
496 TACAGGCTCTATAGATTATGAA 470

seq_name: gb_gss4:AQ705201

seq_documentation_block:

LOCUS AQ705201 453 bp DNA GSS 07-JUL-1999
DEFINITION HS_5522_B2_B06_77A RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate-1098 Col-12 Row-D, genomic survey sequence.
ACCESSION AQ705201
VERSION AQ705201.1 GI:5414627
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

1 (bases 1 to 453)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887

Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.bufo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.bufo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
 Plate: 1098 row: D column: 12
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 453.
 Location/Qualifiers
 source
 1. .453

FEATURES

source
 1. .453
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate:1098 Col=12 Row=D"
 /clone_lib="RPCI-11 Human Male BAC Library"
 /sex="male"
 /note="Vector: pBACe3.6; Genomic sequence of BAC ends"
 BASE COUNT 171 a 71 c 66 g 143 t 2 others
 ORIGIN

alignment_scores:
 Quality: 36.00 Length: 10
 Ratio: 4.000 Gaps: 0
 Percent Similarity: 90.000 Percent Identity: 60.000

alignment_block:

US-08-653-294-9 x AQ705201/rev ..
 Align seg 1/1 to reverse of: AQ705201 from: 1 to: 453

1 TyrArgLeuAlaIleArgLeuAsnGluArg 10
 79 TATTCTATTGCTATGAGATAATGAACGT 50

seq_name: gb_est11:AA270757

seq_documentation_block:
 LOCUS AA270757 555 bp mRNA EST 26-MAR-1997
 DEFINITION va67e06.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone
 IMAGE:736450 5', mRNA sequence.
 ACCESSION AA270757
 VERSION AA270757.1 GI:1909105
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 555)
 REFERENCE
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1398092.
 Contact: Maria M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LBNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:453498
 Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 494.
 Location/Qualifiers
 1. .555

FEATURES

source

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:736450"
 /clone_lib="Soares mouse 3NME12 5"
 /sex="unknown"
 /tissue_type="fetus"
 /dev_stage="12.5dpc total fetus"
 /lab_host="DH10B"
 /note="Organ: whole fetus; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCTATTATTTTTTTTTTTT 3'], on total mouse RNA [provided by Minoru Ko, Wayne State Univ.]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 108 a 191 c 120 g 136 t
 ORIGIN

alignment_scores:
 Quality: 36.00 Length: 10
 Ratio: 3.600 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 60.000

alignment_block:

US-08-653-294-9 x AA270757/rev ..
 Align seg 1/1 to reverse of: AA270757 from: 1 to: 555

1 TyrArgLeuAlaIleArgLeuAsnGluArg 10
 290 TTCGAGTGGCTTAAGACTGAATAACGC 261

seq_name: gb_gss4:AQ735952

seq_documentation_block:
 LOCUS AQ735952 578 bp DNA GSS 15-JUL-1999
 DEFINITION HS_2261.AL_F09_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2261 Col=17 Row=K, genomic survey sequence.
 ACCESSION AQ735952
 VERSION AQ735952.1 GI:5507504
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Mahaliras,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 COMMENT Contact: Mahaliras GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones may be purchased from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
 Plate: 2261 row: K column: 17
 Seq primer: r7
 Class: BAC ends
 High quality sequence stop: 578.
 Location/Qualifiers

```

source
1. .578
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"
BASE COUNT      207 a      97 c      90 g      180 t      4 others
ORIGIN

alignment_scores:
  Quality:      36.00      Length:      10
  Ratio:        4.000      Gaps:        0
  Percent Similarity: 90.000      Percent Identity: 60.000

alignment_block:
US-08-653-294-9 x A0735952 ..
Align seg 1/1 to: A0735952 from: 1 to: 578

1 TyrArgLeuAlaIleArgLeuAsnGluArg 10
|||||
94 TACAGATTATCATGAGATTATGACAGA 123
|||||

seq_name: gb_gss15:A0591361

seq_documentation_block:
LOCUS      A0591361      584 bp      DNA      GSS      08-JUN-1999
DEFINITION HS_5410_B2_E01_T7A_RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate-986 Col-2 Row-J, genomic survey sequence.
ACCESSION  A0591361
VERSION    A0591361.1 GI:5023013
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 584)
AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL    Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE    99380589
COMMENT    On Sep 10, 1998 this sequence version replaced gi:3556571.
            Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 986 row: J column: 2
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 584.
            Location/Qualifiers
                1. .584
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="Plate-986 Col-2 Row-J"
                /clone_lib="RPCI-11 Human Male BAC Library"
                /sex="male"
                /note="Vector: pBACe3.6; Genomic sequence of BAC ends"

FEATURES
            source
1. .578
/organism="Homo sapiens"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="IMAGE:736450"
/clone_lib="Soares mouse 3NME12 5"
/sex="unknown"
/tissue_type="fetus"
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"
/note="Organ: whole fetus; Vector: pT73b-Pac (Pharmacia)
            with a modified polylinker; Site:1: Not I; Site:2: Eco RI;
            1st strand cDNA was primed with a Not I - oligo(dT) primer
            [5' TGTCACCAATCTGAGTGGCGCGCGCTATTTTTTTTTTTT
            3'], on total mouse RNA [provided by Minoru Ko, Wayne
            State Univ.]; double-stranded cDNA was ligated to Eco RI
            adaptors (Pharmacia), digested with Not I and cloned into
            the Not I and Eco RI sites of the modified pT73 vector.

```

```

BASE COUNT      147 a      132 c      103 g      189 t      13 others
ORIGIN

alignment_scores:
  Quality:      36.00      Length:      10
  Ratio:        4.000      Gaps:        0
  Percent Similarity: 90.000      Percent Identity: 70.000

alignment_block:
US-08-653-294-9 x A0591361/rev ..
Align seg 1/1 to reverse of: A0591361 from: 1 to: 584

1 TyrArgLeuAlaIleArgLeuAsnGluArg 10
|||||
232 TATAGACTAGCAATAAATGTCATGAAGA 203
|||||

seq_name: gb_est27:A1463881

seq_documentation_block:
LOCUS      A1463881      585 bp      mRNA      EST      09-MAR-1999
DEFINITION vag7e06.y1 Soares mouse 3NME12 5 Mus musculus cDNA clone IMAGE:736450 5', mRNA sequence.
ACCESSION  A1463881
VERSION    A1463881.1 GI:4317911
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 585)
AUTHORS   Marra,N., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,F.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
            The WashU-NCI Mouse EST Project 1999
            Unpublished (1999)
            On Apr 7, 1998 this sequence version replaced gi:3034529.
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:453498
            This read is a RESEQUENCE of a previously sequenced mouse clone
            correct orientation
            Seq primer: -40RP from Gibco
            High quality sequence stop: 465.
            Location/Qualifiers
                1. 585
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone_lib="IMAGE:736450"
                /clone_lib="Soares mouse 3NME12 5"
                /sex="unknown"
                /tissue_type="fetus"
                /dev_stage="12.5dpc total fetus"
                /lab_host="DH10B"
                /note="Organ: whole fetus; Vector: pT73b-Pac (Pharmacia)
                    with a modified polylinker; Site:1: Not I; Site:2: Eco RI;
                    1st strand cDNA was primed with a Not I - oligo(dT) primer
                    [5' TGTCACCAATCTGAGTGGCGCGCGCTATTTTTTTTTTTT
                    3'], on total mouse RNA [provided by Minoru Ko, Wayne
                    State Univ.]; double-stranded cDNA was ligated to Eco RI
                    adaptors (Pharmacia), digested with Not I and cloned into
                    the Not I and Eco RI sites of the modified pT73 vector.

```


Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 115 a 199 c 131 g 139 t 1 others

alignment_scores:
Quality: 36.00 Length: 10
Ratio: 3.600 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 60.000

alignment_block:

US-08-653-294-9 x AI463881/rev ..

Align seg 1/1 to reverse of: AI463881 from: 1 to: 585

1 TyrArgLeuAlaIleArgLeuAsnGluArg 10

|||||:|||||:|||||:|||||:|||||

290 TTCCGAGTGGCTCTAGACTGAATAACGC 261

seq_name: gb_est32:AI728666

seq_documentation_block:

LOCUS AI728666 626 bp mRNA 11-JUN-1999
DEFINITION BNLGH11358 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (AF032448) ethylene receptor [Malus domestica], mRNA sequence.

ACCESSION AI728666

VERSION AI728666.1 GI:5047518

KEYWORDS EST.

SOURCE upland cotton.

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.

REFERENCE 1 (bases 1 to 626)

ESTs from developing cotton fiber

Blewitt, M., Matz, E.C., Davy, D.F. and Burr, B.

Unpublished (1995)

On Jun 5, 1998 this sequence version replaced gi:3187861.

Contact: Ben Burr

Biology Department

Brookhaven National Laboratory

Upton, NY 11973, USA

Tel: 516-344-3396

Fax: 516-344-3407

Email: burrbnl@bnl.gov

Seq primer: T3 Primer.

Location/Qualifiers

FEATURES

source
1..626
/organism="Gossypium hirsutum"
/cultivar="Acala Maxxa"
/db_xref="taxon:3635"
/clone_lib="Six-day Cotton fiber"
/tissue_type="immature fiber"
/dev_stage="six days post anthesis"
/lab_host="XLI-Blue"
/note="Vector: pBluescript II KS+"

BASE COUNT 177 a 112 c 163 g 174 t

ORIGIN

alignment_scores:
Quality: 36.00 Length: 10
Ratio: 3.600 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 60.000

alignment_block:

US-08-653-294-9 x AI728666 ..

Align seg 1/1 to: AI728666 from: 1 to: 626

1 TyrArgLeuAlaIleArgLeuAsnGluArg 10

|||||:|||||:|||||:|||||:|||||

409 TAGGAAGTCGAATCCGCATAACAGAGA 438

seq_name: gb_gss3:B12156

seq_documentation_block:

LOCUS B12156 1224 bp DNA 14-MAY-1997
DEFINITION F5L16-Sp6 IGF Arabidopsis thaliana genomic clone F5L16, genomic survey sequence.

ACCESSION B12156

VERSION B12156.1 GI:2093276

KEYWORDS GSS.

SOURCE

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 1224)

AUTHORS Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and Ecker, J.

TITLE BAC End Sequences at ATGC

JOURNAL Unpublished (1997)

COMMENT Other GSS: F5L16-T7

Contact: Ecker, J.

Arabidopsis Thaliana Genome Center

University of Pennsylvania

Dept. of Biology, University of Pennsylvania, Philadelphia, PA

19104

Tel: 215-898-9384

Fax: 215-898-8780

Email: jecker@atgenome.bio.upenn.edu

Seq primer: Sp6

Class: BAC ends

High quality sequence start: 80

High quality sequence stop: 762.

FEATURES

source
1..1224
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="F5L16"
/clone_lib="IGF"
/sex="hermaphrodite"
/note="Vector: BelOBAcII; Site_1: EcoRI; Site_2: EcoRI; Produced by Thomas Altmann"

BASE COUNT 397 a 282 c 233 g 303 t

ORIGIN

alignment_scores:
Quality: 36.00 Length: 10
Ratio: 3.600 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 60.000

alignment_block:

US-08-653-294-9 x B12156 ..

Align seg 1/1 to: B12156 from: 1 to: 1224

1 TyrArgLeuAlaIleArgLeuAsnGluArg 10

|||||:|||||:|||||:|||||:|||||

750 TTTCGGCTAGCAATAAAGCTGAACACAGAAA 779

seq_name: gb_est40:AV242923

seq_documentation_block:

LOCUS AV242923 246 bp mRNA EST 04-NOV-1999
DEFINITION AV242923 RIKEN full-length enriched, 0 day neonate head Mus musculus cDNA clone 4831414O15 3', similar to NM_004576 Homo sapiens protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), beta isoform (PPP2R2B) mRNA, mRNA, mRNA sequence.

ACCESSION AV242923

VERSION AV242923.1 GI:6230332

KEYWORDS
SOURCE
ORGANISMEST.
house mouse.
Mus musculusREFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 246)
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T.,
Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F.,
Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I.,
Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M.,
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K.,
Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y.,
Suzuki, H., Suzuki, H., Takahashi, F., Tateno, M., Tomimaga, N.,
Tsunoda, Y., Watanabe, S., Yamamura, T., Yasunishi, A.,
Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Konno, H., et al.)
Unpublished (1999)

TITLE
JOURNAL
COMMENT

On May 18, 1998 this sequence version replaced gi:3137751.
Contact: Yoshihide Hayashizaki
Genome Exploration Research Group, Life Science Tsukuba Center,
Genome Science Laboratory
The Institute of Physical and Chemical Research (RIKEN), Genomic
Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: +81-298-36-9013
Fax: +81-298-36-9098
Email: genome-res@rtc.riken.go.jp/
URL: http://genome.rtc.riken.go.jp/
Sasaki, N., Izawa, M., Watanabe, M., Ozawa, K., Tanaka, T., Yoneda, Y.,
Matsuura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and
Hayashizaki, Y.
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
Itoh, M., Kitsumai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,
Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for
further details.

FEATURES
source

Location/Qualifiers
1..246
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="483141015"
/clone_lib="RIKEN full-length enriched, 0 day neonate
head"
/sex="mixed"
/tissue_type="head"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/note-Site_1: Sali; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 100.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATCTCGAGTTAATTAATATCCCCCCCCCCC
3']. cDNA was cloned into the xhoI and BamHI sites.
Vector: a modified pBluescript KS(+) after bulk excision
from Lambda FLC I."

BASE COUNT 71 a 39 c 37 g 99 t
ORIGIN

alignment_scores:

Quality: 35.00 Length: 9
Ratio: 4.375 Gaps: 0
Percent Similarity: 88.889 Percent Identity: 77.778

alignment_block:

US-08-653-294-9 x AV242923 ..

Align seg 1/1 to: AV242923 from: 1 to: 246

1 TyrArgLeuAlaIleArgLeuAsnGlu 9
|||||:|||||
140 TATAGGTTTCTCATCCGATTAATGAG 166

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 11:54:22 ; Search time 117.7 Seconds
(without alignments)
4.008 Million cell updates/sec

Title: US-08-653-294-10
Perfect score: 49
Sequence: 1 RNLRLRLRY 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : PIR_62: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----------|--------------------|
| 1 | 44 | 89.8 | 137 | 2 I80174 | class I histocompa |
| 2 | 44 | 89.8 | 273 | 2 I38509 | MHC class I histoc |
| 3 | 44 | 89.8 | 274 | 2 I54463 | MHC HLA-B38 chain |
| 4 | 44 | 89.8 | 354 | 2 I59308 | class I histocompa |
| 5 | 44 | 89.8 | 354 | 2 I80168 | class I histocompa |
| 6 | 44 | 89.8 | 354 | 2 I80167 | class I histocompa |
| 7 | 44 | 89.8 | 355 | 2 I80169 | class I histocompa |
| 8 | 44 | 89.8 | 355 | 2 I80171 | class I histocompa |
| 9 | 44 | 89.8 | 359 | 1 HLH012 | MHC class I histoc |
| 10 | 44 | 89.8 | 362 | 1 HLH088 | MHC class I histoc |
| 11 | 44 | 89.8 | 362 | 2 B30345 | MHC class I histoc |
| 12 | 44 | 89.8 | 362 | 2 JH0841 | class I histocompa |
| 13 | 44 | 89.8 | 362 | 2 JH0539 | class I histocompa |
| 14 | 44 | 89.8 | 362 | 2 JH0540 | class I histocompa |
| 15 | 44 | 89.8 | 362 | 2 A45834 | MHC class I histoc |
| 16 | 44 | 89.8 | 362 | 2 I84486 | transmembrane glyc |
| 17 | 44 | 89.8 | 362 | 2 I62045 | gene HLA B-1517 pr |
| 18 | 44 | 89.8 | 362 | 2 I84490 | lymphocyte antigen |
| 19 | 44 | 89.8 | 362 | 2 I37521 | HLA-B*57.2 antigen |
| 20 | 44 | 89.8 | 362 | 2 A30345 | MHC class I histoc |
| 21 | 44 | 89.8 | 362 | 2 I59633 | MHC HLA-B transmem |
| 22 | 44 | 89.8 | 362 | 2 S24434 | class I histocompa |
| 23 | 44 | 89.8 | 362 | 2 I37120 | MHC class I histoc |
| 24 | 44 | 89.8 | 363 | 2 S07113 | class I histocompa |
| 25 | 44 | 89.8 | 363 | 2 S03537 | class I histocompa |
| 26 | 44 | 89.8 | 364 | 2 D35997 | MHC class I histoc |
| 27 | 44 | 89.8 | 365 | 2 S77963 | MHC class I histoc |
| 28 | 44 | 89.8 | 365 | 2 I54416 | HLA-AW24 protein - |
| 29 | 44 | 89.8 | 365 | 2 I54493 | MHC class I histoc |
| 30 | 39 | 79.6 | 274 | 1 HLH032 | MHC class I histoc |

31 39 79.6 338 2 I56116 MHC HLA-B27-HS - h
32 39 79.6 335 2 I37516 HLA-B alpha-chain
33 39 79.6 362 1 HLH082 MHC class I histoc
34 39 79.6 362 2 C35997 MHC class I histoc
35 39 79.6 362 2 S25415 class I histocompa
36 39 79.6 362 2 A45850 MHC class I histoc
37 39 79.6 362 2 I68724 MHC class I histoc
38 39 79.6 362 2 I61861 MHC HLA-B44.2 chal
39 39 79.6 362 2 I37485 human lymphocyte a
40 39 79.6 362 2 I54442 MHC class I histoc
41 39 79.6 362 2 I37515 MHC class I histoc
42 39 79.6 362 2 I54289 MHC HLA-B27d - hum
43 39 79.6 364 2 A35997 MHC class I histoc
44 38 77.6 362 2 I54457 MHC class I lympho
45 37 75.5 218 2 I72808 MHC class I HLA-J

ALIGNMENTS

RESULT 1

I80174
class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999
C:Accession: I80174
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Wat
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I80174
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-137 <RES>
A:Cross-references: EMBL:U05585; NID:g454787; PIDN:AAA50188.1; PID:g454788
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 89.8%; Score 44; DB 2; Length 137;
Best Local Similarity 90.0%; Pred. No. 0.074;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RNLRLRLRY 10
DB 40 RNLRLRLRY 49

RESULT 2

I38509
MHC class I histocompatibility antigen - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 23-Jul-1999
C:Accession: I38509
R:Cereb, N.; Choi, J.W.; Riu, K.Z.; Yang, S.Y.
Tissue Antigens 44, 271-273, 1994
A:Title: HLA-B*5105, a newly identified B51 IEF variant.
A:Reference number: I38509; MUID:95176331
A:Accession: I38509
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: mRNA
A:Residues: 1-273 <RES>
A:Cross-references: EMBL:U06697; NID:g469544; PIDN:AAA92997.1; PID:g469545
C:Genetics:
A:Gene: GDB:HLA-B
A:Cross-references: GDB:120048; OMIM:142830
A:Map position: 6p21.3-6p21.3
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 89.8%; Score 44; DB 2; Length 273;
Best Local Similarity 90.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENRILLRY 10
||||| |||
Db 74 RENRILRY 83

RESULT 3
MHC HLA-B38 chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
C:Accession: I54463
R:McAdam, S.N.; Engler-Blum, G.; Gekeler, V.; Steiert, I.; Weiss, E.; Schmidt, H.
Immunogenetics 30, 200-207, 1989
A:Title: Genetic and serological heterogeneity of the supertypic HLA-B locus specificities
A:Reference number: I54463; MUID:89379286
A:Accession: I54463
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-274 <RES>
A:Cross-references: GB:M29864; NID:g187674; PIDN:AAA36222.1; PID:g187675
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 89.8%; Score 44; DB 2; Length 274;
Best Local Similarity 90.0%; Pred. No. 0.15; Mismatches 1; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 RENRILLRY 10
||||| |||
Db 75 RENRILRY 84

RESULT 4
class I histocompatibility antigen - pygmy chimpanzee (fragment)
C:Species: Pan paniscus (pygmy chimpanzee, bonobo)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 23-Jul-1999
C:Accession: I59308
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I59308
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-354 <RES>
A:Cross-references: EMBL:U05575; NID:g454767; PIDN:AAA50178.1; PID:g454768
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 89.8%; Score 44; DB 2; Length 354;
Best Local Similarity 90.0%; Pred. No. 0.2; Mismatches 1; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 RENRILLRY 10
||||| |||
Db 91 RENRILRY 100

RESULT 5
class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999
C:Accession: I80168
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I80168
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-354 <RES>

A:Cross-references: EMBL:U05579; NID:g454775; PIDN:AAA50182.1; PID:g454776
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 89.8%; Score 44; DB 2; Length 354;
Best Local Similarity 90.0%; Pred. No. 0.2; Mismatches 1; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 RENRILLRY 10
||||| |||
Db 91 RENRILRY 100

RESULT 6
class I histocompatibility antigen - pygmy chimpanzee (fragment)
C:Species: Pan paniscus (pygmy chimpanzee, bonobo)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999
C:Accession: I80167
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Wat
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I80167
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-354 <RES>
A:Cross-references: EMBL:U05578; NID:g454773; PIDN:AAA50181.1; PID:g454774
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 89.8%; Score 44; DB 2; Length 354;
Best Local Similarity 90.0%; Pred. No. 0.2; Mismatches 1; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 RENRILLRY 10
||||| |||
Db 91 RENRILRY 100

RESULT 7
class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999
C:Accession: I80169
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Wat
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544
A:Accession: I80169
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-355 <RES>
A:Cross-references: EMBL:U05580; NID:g454777; PIDN:AAA50183.1; PID:g454778
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 89.8%; Score 44; DB 2; Length 355;
Best Local Similarity 90.0%; Pred. No. 0.2; Mismatches 1; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 RENRILLRY 10
||||| |||
Db 91 RENRILRY 100

RESULT 8
class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999
C:Accession: I80171

R;McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: 159308; MUID:94286544
A:Accession: I80171
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-355 <RES>
A:Cross-references: EMBL:U05582; NID:9454781; PIDN:AAA50185.1; PID:9454782
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 89.8%; Score 44; DB 2; Length 355;
Best Local Similarity 90.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENRILRLRY 10
||||| |||

Db 91 RENRILRLRY 100

RESULT 9

HLH012
MHC class I histocompatibility antigen HLA alpha chain precursor (clone pHLA 12.4) - hum
C:Species: Homo sapiens (man)
C:Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 22-Jun-1999
C:Accession: A02189
R;Malissen, M.; Mallissen, B.; Jordan, B.R.
Proc. Natl. Acad. Sci. U.S.A. 79, 893-897, 1982
A:Title: Exon/intron organization and complete nucleotide sequence of an HLA gene.
A:Reference number: A02189; MUID:82151002
A:Accession: A02189
A:Molecule type: DNA
A:Residues: 1-359 <MAL>
A:Cross-references: GB:J00191; GB:V00526; NID:q187600; PIDN:AAA36218.1; PID:G386873
C:Comment: The seven exons correspond approximately to the domain structure of this chain
C:Genetics:
A:Map position: 6p21.3
A:Introns: 22/1; 112/1; 204/1; 296/1; 335/1; 346/1
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplanted
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-359/Product: class I histocompatibility antigen HLA alpha chain #status predicted <
F:22-304/Domain: extracellular #status predicted <EXT>
F:22-111/Domain: alpha-1 <EX1>
F:112-203/Domain: immunoglobulin homology <IMM>
F:217-282/Domain: transmembrane #status predicted <TM>
F:305-329/Domain: transmembrane #status predicted <TM>
F:335-359/Domain: intracellular #status predicted <INT>
F:107/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:224-280/Disulfide bonds: #status predicted

Query Match 89.8%; Score 44; DB 1; Length 359;
Best Local Similarity 90.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENRILRLRY 10
||||| |||

Db 96 RENRILRLRY 105

RESULT 10

HLH08
MHC class I histocompatibility antigen HLA-Bw58 alpha chain precursor - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 05-Sep-1997
C:Accession: A23895
R;Ways, J.P.; Coppin, H.L.; Parham, P.
J. Biol. Chem. 260, 11924-11933, 1985
A:Title: The complete primary structure of HLA-Bw58.
A:Reference number: A23895; MUID:86008247
A:Accession: A23895
A:Cross-references: EMBL:X60254; NID:g22869; PIDN:CAA42806.1; PID:g22870

A:Molecule type: DNA
A:Residues: 1-362 <WAY>
A:Note: the authors translated the codon GCC for residue 349 as Ser
C:Comment: This protein is a subtype of the HLA-B*7 family.
C:Genetics:
A:Gene: GDB:HLA-B
A:Cross-references: GDB:120048; OMIM:142830
A:Map position: 6p21.3-6p21.3
A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplant
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-362/Product: class I histocompatibility antigen HLA-Bw58 alpha chain #status pre
F:25-307/Domain: extracellular #status predicted <EXT>
F:25-114/Domain: alpha-1 <EX1>
F:115-206/Domain: alpha-2 <EX2>
F:220-285/Domain: immunoglobulin homology <IMM>
F:308-331/Domain: transmembrane #status predicted <TM>
F:332-362/Domain: intracellular #status predicted <INT>
F:110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 89.8%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENRILRLRY 10
||||| |||

Db 99 RENRILRLRY 108

RESULT 11

B30345
MHC class I histocompatibility antigen HLA-Bw52 precursor - human
C:Species: Homo sapiens (man)
C:Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 16-Feb-1997
C:Accession: B30345
R;Hayashi, H.; Ennis, P.D.; Ariga, H.; Salter, R.D.; Parham, P.; Kano, K.; Takiguchi, J.
Immunol. 142, 306-311, 1989
A:Title: HLA-B*51 and HLA-Bw52 differ by only two amino acids which are in the helical
A:Reference number: A30345; MUID:89080265
A:Accession: B30345
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-362 <HAY>
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: transmembrane protein
F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 89.8%; Score 44; DB 2; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENRILRLRY 10
||||| |||

Db 99 RENRILRLRY 108

RESULT 12

JH0541
class I histocompatibility antigen Gogo-B0103 heavy chain precursor - lowland gorilla
C:Species: Gorilla gorilla gorilla (lowland gorilla)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C:Accession: JH0541
R;Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.
J. Exp. Med. 174, 1491-1509, 1991
A:Title: Gorilla class I major histocompatibility complex alleles: comparison to huma
A:Reference number: JH0534; MUID:92078860
A:Accession: JH0541
A:Molecule type: DNA
A:Residues: 1-362 <LAW>
A:Cross-references: EMBL:X60254; NID:g22869; PIDN:CAA42806.1; PID:g22870

A:Experimental source: EBV-transformed B cell

C:Genetics:

A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

C:Keywords: transmembrane protein

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0103 #status pre

F:25-114/Domain: alpha-1 <AL1>

F:115-206/Domain: alpha-2 <AL2>

F:207-298/Domain: alpha-3 <AL3>

F:220-285/Domain: immunoglobulin homology <IMM>

F:299-362/Domain: intracellular #status predicted <INT>

Query Match 89.8%; Score 44; DB 2; Length 362;

Best Local Similarity 90.0%; Pred. No. 0.21;

Matches 9; Conservative 0; Mismatches 0; Gaps 0; Indels 1;

QY 1 RENLRILRY 10

||||| |||

Db 99 RENLRILRY 108

RESULT 13

JH0539

class I histocompatibility antigen Gogo-B0101 heavy chain precursor - lowland gorilla

C:Species: Gorilla gorilla gorilla (lowland gorilla)

C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999

C:Accession: JH0539

R:Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.

J. Exp. Med. 174, 1491-1509, 1991

A:Title: Gorilla class I major histocompatibility complex alleles: comparison to human a

A:Reference number: JH0534; MUID:92078860

A:Accession: JH0539

A:Molecule type: DNA

A:Residues: 1-362 <LAW>

A:Cross-references: EMBL:X60255; NID:g22865; PIDN:CAA42807.1; PID:g22866

A:Experimental source: EBV-transformed B cell

C:Genetics:

A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

C:Keywords: transmembrane protein

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0101 #status pre

F:25-114/Domain: alpha-1 <AL1>

F:115-206/Domain: alpha-2 <AL2>

F:207-298/Domain: alpha-3 <AL3>

F:220-285/Domain: immunoglobulin homology <IMM>

F:299-362/Domain: intracellular #status predicted <INT>

Query Match 89.8%; Score 44; DB 2; Length 362;

Best Local Similarity 90.0%; Pred. No. 0.21;

Matches 9; Conservative 0; Mismatches 0; Gaps 0; Indels 1;

QY 1 RENLRILRY 10

||||| |||

Db 99 RENLRILRY 108

RESULT 14

JH0540

class I histocompatibility antigen Gogo-B0102 heavy chain precursor - lowland gorilla

C:Species: Gorilla gorilla gorilla (lowland gorilla)

C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999

C:Accession: JH0540

R:Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.

J. Exp. Med. 174, 1491-1509, 1991

A:Title: Gorilla class I major histocompatibility complex alleles: comparison to human a

A:Reference number: JH0534; MUID:92078860

A:Accession: JH0540

A:Molecule type: DNA

A:Residues: 1-362 <LAW>

A:Cross-references: EMBL:X60693; NID:g22867; PIDN:CAA43101.1; PID:g22868

A:Experimental source: EBV-transformed B cell

C:Genetics:

A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

C:Keywords: transmembrane protein

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0102 #status

F:25-114/Domain: alpha-1 <AL1>

F:115-206/Domain: alpha-2 <AL2>

F:207-298/Domain: alpha-3 <AL3>

F:220-285/Domain: immunoglobulin homology <IMM>

F:299-362/Domain: intracellular #status predicted <INT>

Query Match 89.8%; Score 44; DB 2; Length 362;

Best Local Similarity 90.0%; Pred. No. 0.21;

Matches 9; Conservative 0; Mismatches 0; Gaps 0; Indels 1;

QY 1 RENLRILRY 10

||||| |||

Db 99 RENLRILRY 108

RESULT 15

A45834

MHC class I histocompatibility antigen HLA-B53 alpha chain precursor - human

C:Species: Homo sapiens (man)

C>Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 23-Jul-1999

C:Accession: A45834

R:Hayashi, H.; Ooba, T.; Nakayama, S.; Sekimata, M.; Kano, K.; Takiguchi, M.

Immunogenetics 32, 195-199, 1990

A:Title: Allospecificities between HLA-Bw53 and HLA-B35 are generated by substitution

A:Reference number: A45834; MUID:91033941

A:Accession: A45834

A:Molecule type: DNA

A:Residues: 1-362 <HAY>

A:Cross-references: GB:M58636; NID:gl87756; PIDN:AAA36228.1; PID:gl87757; GB:M33574

A:Note: this allele is designated B*5301

C:Genetics:

A:Gene: GDB:HLA-B

A:Cross-references: GDB:120048; OMIM:142830

A:Map position: 6p21.3-6p21.3

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

C:Keywords: glycoprotein; heterodimer; transmembrane protein

F:1-24/Domain: signal sequence #status predicted <SIG>

F:220-285/Domain: immunoglobulin homology <IMM>

F:110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 89.8%; Score 44; DB 2; Length 362;

Best Local Similarity 90.0%; Pred. No. 0.21;

Matches 9; Conservative 0; Mismatches 0; Gaps 0; Indels 1;

QY 1 RENLRILRY 10

||||| |||

Db 99 RENLRILRY 108

Search completed: February 7, 2000, 11:54:22

Job time: 24332 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:29:37 ; Search time 122.56 Seconds
(without alignments)
1.933 Million cell updates/sec

Title: US-08-653-294-10
Perfect score: 49
Sequence: 1 RENRILLRY 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|-------|-------|--------|----------|----------------------|
| 1 | 49 | 100.0 | 10 | 1 R83096 | HLA-B2702 CTL modu |
| 2 | 49 | 100.0 | 10 | 1 W47267 | Immunomodulatory p |
| 3 | 49 | 100.0 | 10 | 1 W3789 | Peptide B2702.75-8 |
| 4 | 44 | 89.8 | 10 | 1 R41208 | Peptide fragment o |
| 5 | 44 | 89.8 | 10 | 1 R83062 | HLA-B2702 CTL modu |
| 6 | 44 | 89.8 | 10 | 1 R95413 | Alpha1-helix of HL |
| 7 | 44 | 89.8 | 10 | 1 R95427 | HLA-B2702.75-84(L) |
| 8 | 44 | 89.8 | 10 | 1 W07512 | T-cell modulating |
| 9 | 44 | 89.8 | 10 | 1 W07514 | T-cell modulating |
| 10 | 44 | 89.8 | 10 | 1 W47265 | Immunomodulatory p |
| 11 | 44 | 89.8 | 10 | 1 W47271 | Immunomodulatory p |
| 12 | 44 | 89.8 | 10 | 1 W3784 | Peptide B2702.75-8 |
| 13 | 44 | 89.8 | 15 | 1 R92912 | HLA-B2702 CTL modu |
| 14 | 44 | 89.8 | 15 | 1 W3795 | Peptide B2702.70-8 |
| 15 | 44 | 89.8 | 20 | 1 R92907 | HLA-B2702 CTL modu |
| 16 | 44 | 89.8 | 20 | 1 R92908 | HLA-B2702 CTL modu |
| 17 | 44 | 89.8 | 20 | 1 R95428 | HLA-B2702 84-75-84 |
| 18 | 44 | 89.8 | 20 | 1 W3778 | Immunomodulating d |
| 19 | 44 | 89.8 | 20 | 1 W3791 | Peptide B2702.84-7 |
| 20 | 44 | 89.8 | 25 | 1 R41205 | Peptide fragment o |
| 21 | 44 | 89.8 | 25 | 1 R48286 | Peptide fragment o |
| 22 | 44 | 89.8 | 25 | 1 R83090 | HLA-B2702 CTL modu |
| 23 | 44 | 89.8 | 25 | 1 R83093 | HLAB38 CTL modulatio |
| 24 | 44 | 89.8 | 25 | 1 R95416 | HLA-B2702.60-84. C |
| 25 | 44 | 89.8 | 25 | 1 R95422 | HLAB38.6084. Comps |
| 26 | 44 | 89.8 | 25 | 1 W3794 | Peptide B2702.60-8 |
| 27 | 44 | 89.8 | 184 | 1 Y06801 | Peptide Seq ID No: |
| 28 | 44 | 89.8 | 362 | 1 R03142 | Sequence of HLA-Bw |
| 29 | 44 | 89.8 | 362 | 1 R03144 | Sequence of HLA-B5 |
| 30 | 44 | 89.8 | 362 | 1 R12463 | HLA-Bw53 exon. HLA |
| 31 | 39 | 79.6 | 10 | 1 R41212 | Peptide fragment o |
| 32 | 39 | 79.6 | 10 | 1 R83075 | HLA-B2702 CTL modu |
| 33 | 39 | 79.6 | 10 | 1 R83094 | HLA-B2702 CTL modu |
| 34 | 39 | 79.6 | 10 | 1 R83095 | HLA-B2702 CTL modu |

35 39 79.6 10 1 R95423 HLA-B2705.75-84. C
36 39 79.6 10 1 R95425 HLA-B2702.75-84 (D)
37 39 79.6 10 1 R95426 HLA-B2702.75-84 (T)
38 39 79.6 10 1 W07513 T-cell modulating
39 39 79.6 10 1 W47269 Immunomodulatory p
40 39 79.6 10 1 W3785 Peptide B2705.75-8
41 39 79.6 10 1 W3788 Peptide B2702.75-8
42 39 79.6 10 1 W3787 Peptide B2702.75-8
43 39 79.6 17 1 R71442 Human HLA-B27-(62-
44 39 79.6 17 1 R71443 Human [Phe74]-HLA-
45 39 79.6 20 1 R92909 HLA-B2702 CTL modu

ALIGNMENTS

RESULT 1

R83096
ID R83096 standard; peptide; 10 AA.
AC R83096;
DT 16-MAY-1996 (first entry)
DE HLA-B2702 CTL modulating peptide (B2702.75-84(L)).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW Class I MHC; HLA-B2702.
OS Synthetic.
PN W09526979-AL.
PD 12-OCT-1995.
PF 05-APR-1995; U04349.
PR 05-APR-1994; US-222851.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM, Parham P;
DR WPI; 95-358582/46.
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
PS Example 14; Page 34; 80pp; English.
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of
CC class I major histocompatibility complex (MHC) antigens. This sequence
CC corresponds to residues 75-84 of the alpha-1 domain of the class I MHC
CC HLA-B2702. These sequences can be used to extend the period of
CC acceptance by a recipient of a transplant from an MHC unmatched donor.
CC The peptides are administered to a patient in conjunction with
CC a subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.
SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00049;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENRILLRY 10
| | | | |
DB 1 RENRILLRY 10

RESULT 2

W47267
ID W47267 standard; peptide; 10 AA.
AC W47267;
DT 22-MAY-1998 (first entry)
DE Immunomodulatory peptide.
KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
KW transplant rejection; treatment; autoimmune disease.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 1..10 /note= "at least one of the amino acids is the

FT WO9744052-A1.
 PN D-isomer
 PD 27-NOV-1997.
 PF 23-APR-1997; U06705.
 PR 22-MAY-1996; US-651650.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI: 98-018220/02.
 PT Novel immunomodulatory peptide-type compound - useful for inhibiting
 PT transplant rejection
 PS Claim 10: Page 36; 41pp; English.
 CC The present sequence is an immunomodulatory peptide, which
 CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
 CC in a pharmaceutical composition together with a subtherapeutic dose
 CC of an immunosuppressant, to extend the period of acceptance of a
 CC transplant from a major histocompatibility complex (MHC) unmatched
 CC donor, i.e. to inhibit transplant rejection. It can also be used in
 CC the treatment of autoimmune diseases.
 CC Peptides using the D-form amino acids are more effective
 CC immunomodulators than their diastereomers or enantiomers.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.00049; Mismatches 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0;

QY 1 RENRLILRY 10
 DB 1 RENRLILRY 10

RESULT 3
 ID W33789 standard; peptide; 10 AA.
 AC W33789;
 DT 19-JUN-1998 (first entry)
 DE Peptide B2702.75-84L81 tested for immunomodulating activity.
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN WO9744351-A1.
 PD 27-NOV-1997.
 PF 22-MAY-1997; U08689.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI: 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Example 1; Page 19; 41pp; English.
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
 CC activity. A peptide-type compound or variant is claimed which has
 CC immunomodulating activity, including the N-terminal acylated and/or
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where
 CC the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha-1 domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.

SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.00049; Mismatches 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0;

QY 1 RENRLILRY 10
 DB 1 RENRLILRY 10

RESULT 4
 ID R41208 standard; peptide; 10 AA.
 AC R41208;
 DT 15-MAR-1994 (first entry)
 DE Peptide fragment of Class I HLA peptide.
 KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;
 KW parasitic disease; cytotoxic T lymphocyte; modulation.
 OS Synthetic.
 PN WO9317699-A.
 PD 16-SEP-1993.
 PF 25-FEB-1993; U01758.
 PR 02-MAR-1992; US-844716.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger CA, Krensky AM;
 DR WPI: 93-303134/38.
 PT New peptide(s) based on Class I HLA antigen domains - used for
 PT modulating cytotoxic T-lymphocyte activity towards targets
 PS Claim 11; Page 54; 61pp; English.
 CC The peptide is used to modulate cytotoxic T-lymphocyte (CTL)
 CC activity, either by inhibition or stimulation. It can be used
 CC for inhibiting CTL toxicity in transplantations, for inducing CTL
 CC activity in parasitic diseases and neoplasia and in studies on viral
 CC infection. The peptide can also be used for identifying CTLs which
 CC bind to it and removing subsets of CTLs from a T-cell composition.
 CC This peptide sequence is more commonly found within larger peptide
 CC compounds of not more than 30 amino acids in length.
 SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0048; Mismatches 1; Indels 0; Gaps 0;
 Matches 9; Conservative 0;

QY 1 RENRLILRY 10
 DB 1 RENRLILRY 10

RESULT 5
 ID R83062 standard; peptide; 10 AA.
 AC R83062;
 DT 16-MAY-1996 (first entry)
 DE HLA-B2702 CTL modulating peptide (B2702.75-84).
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW class I MHC; HLA-B2702.
 OS Synthetic.
 PN WO9528979-A1.
 PD 12-OCT-1995.
 PF 05-APR-1995; U04349.
 PR 05-APR-1994; US-222851.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM, Farham P;
 DR WPI: 95-358582/46.
 PT Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 PT host
 PS Claim 15; Page 9; 80pp; English.
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of

CC class I major histocompatibility complex (MHC) antigens. This sequence
 CC corresponds to residues 75-84 of the alpha-1 domain of the class I MHC
 CC HLA-B*2702. These sequences can be used to extend the period of
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.
 CC The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0048;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENRLRLRY 10
 DB 1 RENRLRLRY 10

RESULT 6

R95413
 ID R95413 standard; peptide; 10 AA.
 AC R95413; 1996 (first entry)
 DE Alpha-helix of HLA-B*2702.
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KW cytolysis; antigen presenting cell.
 OS Synthetic.
 PN WO9513288-A1.
 PD 18-MAY-1995.
 PF 10-NOV-1994; US-150493.
 PR 10-NOV-1993; US-150493.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI: 95-194027/25.
 PT Compsns. comprising lymphoid surface membrane proteins - which may
 PT inhibit cytolytic activity and differentiation of CTLs.
 PS Example; Page 11; 29pp; English.

CC This sequence represents the alpha-helix of the
 CC human-leucocyte-associated antigen B2702 (HLA-B*2702). This sequence,
 CC epitopes, and palindromes of it (such as R95428) can be used to isolate
 CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
 CC protein associated with T-cell activation in mammalian T-cells, and is
 CC also immunologically cross reactive with the heat shock protein Hsc70.
 CC p74 is found in a limited number of cell types, but is particularly
 CC expressed on B and T cells. p74 can be isolated by lysis of a suitable
 CC cell with an amphoteric detergent, and then passed through an affinity
 CC column containing a covalently bound HLA-B*2702 palindromic peptide.
 CC Compositions comprising the extracellular fragment of p74 combined with
 CC HLA-B*2702.60-84 (see R95416), induces calcium influx, and inhibits
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate
 CC compounds can be screened for their effect on the cytolytic activity of
 CC T-cells, by combining them with the extracellular portion of p74 and
 CC determining the amount of binding between the candidate compound and p74.
 CC Modulation of CTL activity can be inhibited in a cellular composition
 CC containing T-cells and antigen presenting cells (APCs), by adding to the
 CC mix the extracellular portion of p74, in an amount sufficient to compete
 CC with p74 for the binding of the p74 ligand.
 SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0048;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENRLRLRY 10
 DB 1 RENRLRLRY 10

RESULT 7

R95427
 ID R95427 standard; peptide; 10 AA.
 AC R95427;
 DT 12-NOV-1996 (first entry)
 DE HLA-B*2702.75-84(L).
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KW cytolysis; antigen presenting cell.
 OS Synthetic.

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0048;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENRLRLRY 10
 DB 1 RENRLRLRY 10

RESULT 8

R95427
 ID R95427 standard; peptide; 10 AA.
 AC R95427;
 DT 12-NOV-1996 (first entry)
 DE HLA-B*2702.75-84(L).
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KW cytolysis; antigen presenting cell.
 OS Synthetic.
 PN WO9513288-A1.
 PD 18-MAY-1995.
 PF 10-NOV-1994; US-150493.
 PR 10-NOV-1993; US-150493.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI: 95-194027/25.
 PT Compsns. comprising lymphoid surface membrane proteins - which may
 PT inhibit cytolytic activity and differentiation of CTLs.
 PS Example; Page 12; 29pp; English.

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0048;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENRLRLRY 10
 DB 1 RENRLRLRY 10

RESULT 9

R95427
 ID R95427 standard; peptide; 10 AA.
 AC R95427;
 DT 12-NOV-1996 (first entry)
 DE HLA-B*2702.75-84(L).
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KW cytolysis; antigen presenting cell.
 OS Synthetic.
 PN WO9513288-A1.
 PD 18-MAY-1995.

QY 1 RENRLRLRY 10
 DB 1 RENRLRLRY 10

PF 05-APR-1996; U04710.
 PR 12-MAY-1995; US-440504.
 PA (SANG-) SANGSTAT MEDICAL CORP.
 PI Buelow R;
 DR WPI: 96-518410/51.
 PT Treatment of auto-immune disease by admin. of peptide(s) corresp. to
 PT major histocompatibility complex antigens - esp. for delaying onset
 PT of clinical symptoms of insulin dependent diabetes by modulating T
 PT cell mediated attack on target cells
 PS Claim 7; Page 20; 24pp; English.
 CC W07512-W07518 represent T-cell modulating peptides that can be used in
 CC the method of the invention. These sequences are based on a portion of
 CC the generic peptide corresponding to residues 70-91 of the alpha-1 domain
 CC of the major histocompatibility complex (MHC) class I antigen (see
 CC W07510). The method is for affecting the course of an autoimmune disease
 CC involving T-cell mediated destruction of tissue in mammals. These
 CC peptides are used especially to treat insulin-dependent diabetes
 CC mellitus, preferably being administered during the pre-clinical stage to
 CC delay onset of the disease. Other diseases that can be treated are
 CC Sjogren's disease, rheumatoid arthritis, psoriasis, pemphigus vulgaris,
 CC myasthenia gravis, etc. The peptides modulate T-cell mediated attack on autologous
 CC target cells, and may also reduce inflammation, swelling, and release of
 CC cytokines, perforins, granzymes etc. associated with T cell activation.
 CC Sequence 10 AA.
 SQ

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0048;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 RENRILLY 10
 Db 1 RENRILLY 10
 RESULT 10
 W47265
 ID W47265 standard; peptide; 10 AA.
 AC W47265; 1998 (first entry)
 DE Immunomodulatory peptide.
 KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
 KW transplant rejection; treatment; autoimmune disease.
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc_difference 1.10 /note- "at least one of the amino acids is the
 FT D-isomer
 FT
 PN W09744052-A1.
 PD 27-NOV-1997.
 PF 23-APR-1997; U06705.
 PR 22-MAY-1996; US-651650.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI: 98-018220/02.
 PT Novel immunomodulatory peptide-type compound - useful for inhibiting
 PT transplant rejection
 PS Claim 10; Page 36; 41pp; English.
 CC The present sequence is an immunomodulatory peptide, which
 CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
 CC in a pharmaceutical composition together with a subtherapeutic dose
 CC of an immunosuppressant, to extend the period of acceptance of a
 CC transplant from a major histocompatibility complex (MHC) unmatched
 CC donor, i.e. to inhibit transplant rejection. It can also be used in
 CC the treatment of autoimmune diseases.
 CC Peptides using the D-form amino acids are more effective
 CC immunomodulators than their diastereomers or enantiomers.
 CC Sequence 10 AA;
 SQ

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0048;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 RENRILLY 10
 Db 1 RENRILLY 10
 RESULT 11
 W47271
 ID W47271 standard; peptide; 10 AA.
 AC W47271;
 DT 22-MAY-1998 (first entry)
 DE Immunomodulatory peptide.
 KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
 KW transplant rejection; treatment; autoimmune disease.
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc_difference 1.10 /note- "at least one of the amino acids is the
 FT D-isomer
 FT

PN W09744052-A1.
 PD 27-NOV-1997.
 PF 23-APR-1997; U06705.
 PR 22-MAY-1996; US-651650.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI: 98-018220/02.
 PT Novel immunomodulatory peptide-type compound - useful for inhibiting
 PT transplant rejection
 PS Claim 10; Page 36; 41pp; English.
 CC The present sequence is an immunomodulatory peptide, which
 CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
 CC in a pharmaceutical composition together with a subtherapeutic dose
 CC of an immunosuppressant, to extend the period of acceptance of a
 CC transplant from a major histocompatibility complex (MHC) unmatched
 CC donor, i.e. to inhibit transplant rejection. It can also be used in
 CC the treatment of autoimmune diseases.
 CC Peptides using the D-form amino acids are more effective
 CC immunomodulators than their diastereomers or enantiomers.
 SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0048;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RENLRILLRY 10
 ||:|||||
 Db 1 REDLRILLRY 10

RESULT 12

W33784
 ID W33784 standard; peptide; 10 AA.
 AC W33784;
 DT 19-JUN-1998 (first entry)
 DE Peptide B2702.75-84 tested for immunomodulating activity.
 KW Immunomodulating dimer; Immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN W09744351-A1.
 PD 27-NOV-1997.
 PF 22-MAY-1997; U08689.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI: 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Example 1; Page 19; 41pp; English.
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
 CC activity. A peptide-type compound or variant is claimed which has
 CC immunomodulating activity, including the N-terminal acylated and/or
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where
 CC the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha-1 domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.0048;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RENLRILLRY 10
 |||||||
 Db 1 RENLRILLRY 10

RESULT 13

R92912
 ID R92912 standard; peptide; 15 AA.
 AC R92912;
 DT 16-MAY-1996 (first entry)
 DE HLA-B2702 CTL modulating peptide (B2702.70-84).
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW Class I MHC; HLA-B2702.
 OS Synthetic.
 PN W09526979-A1.
 PD 12-OCT-1995.
 PF 05-APR-1995; U04349.
 PR 03-APR-1994; US-222851.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM, Parham P;
 DR WPI: 95-358582/46.
 PT Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B7-84 MHC antigen of the recipient
 PT host
 PS Example 15; Page 36; 80pp; English.
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of
 CC Class I major histocompatibility complex (MHC) antigens. This sequence
 CC corresponds to residues 70-84 of the alpha-1 domain of the class I MHC
 CC HLA-B2702. These sequences can be used to extend the period of
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.
 CC The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 SQ Sequence 15 AA;

Query Match 89.8%; Score 44; DB 1; Length 15;
 Best Local Similarity 90.0%; Pred. No. 0.0075;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

-QY 1 RENLRILLRY 10
 |||||||
 Db 6 RENLRILLRY 15

RESULT 14

W33795
 ID W33795 standard; peptide; 15 AA.
 AC W33795;
 DT 19-JUN-1998 (first entry)
 DE Peptide B2702.70-84 tested for immunomodulating activity.
 KW Immunomodulating dimer; Immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN W09744351-A1.
 PD 27-NOV-1997.
 PF 22-MAY-1997; U08689.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI: 98-086530/08.

PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
PT alpha-1 domain, used for preventing rejection of transplants or
PT treating autoimmune diseases
PS Example 1: Page 19: 41pp; English.
CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
CC activity. A peptide-type compound or variant is claimed which has
CC immunomodulating activity, including the N-terminal acylated and/or
CC C-terminal amidated or esterified forms of up to 60 amino acids, where
CC the peptide-type compound comprises the formula: A-B, where A, B =
CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
CC acid. The sequence in the brackets may optionally be absent or truncated
CC at any peptide type bond within the brackets. The compounds comprise
CC amino acid sequences related to a Class I HLA-B alpha domain (positions
CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
CC undesirably attacking cells in a host or in vitro. They can also be
CC used in combination with antigenic peptides or proteins of interest to
CC activate CTLs. They can also inhibit the proliferation of T cells in
CC response to anti-CD3. The peptide can be used for preventing rejection
CC of transplants or for treating autoimmune diseases, e.g. diabetes,
CC rheumatoid arthritis and lupus erythematosus. The products can also be
CC used for detection and diagnosis.
SQ Sequence 15 AA;

Query Match 89.8%; Score 44; DB 1; Length 15;
Best Local Similarity 90.0%; Pred. No. 0.0075;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRILLY 10
DB 6 RENTRIALRY 15
||||| |||

RESULT 15

R92907
ID R92907 standard; peptide: 20 AA.
AC R92907;
DT 16-MAY-1996 (first entry)
DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW immunosuppressant; graft versus host disorder; transplantation; therapy;
KW class I MHC; HLA-B2702.
OS Synthetic.
PN WO9526979-A1.
PD 12-OCT-1995.
PF 05-APR-1995; U04349.
PR 05-APR-1994; US-222851.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM, Parham P;
DR WPI; 95-358582/46.
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
PS Example 15; Page 36; 80pp; English.
CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of
CC class I major histocompatibility complex (MHC) antigens. This sequence
CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class
CC I MHC HLA-B2702. These sequences can be used to extend the period of
CC acceptance by a recipient of a transplant from an MHC unmatched donor.
CC The peptides are administered to a patient in conjunction with a
CC subtherapeutic amount of an immunosuppressant. This is administered to
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.
SQ Sequence 20 AA;

Query Match 89.8%; Score 44; DB 1; Length 20;
Best Local Similarity 90.0%; Pred. No. 0.01;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRILLY 10
DB 11 RENTRIALRY 20
||||| |||

Search completed: February 8, 2000, 01:29:37
Job time: 1749 sec

| Result No. | Score | Query | | Length | DB | ID | Description |
|------------|-------|-------|-----|--------|-------------|----|--------------------|
| | | Match | | | | | |
| 1 | 44 | 89.8 | 359 | 1 | 1B01_PANTR | | P13750 pan troglod |
| 2 | 44 | 89.8 | 362 | 1 | 1B01_GORGO | | P33799 gorilla gor |
| 3 | 44 | 89.8 | 362 | 1 | 1B02_GORGO | | P30380 gorilla gor |
| 4 | 44 | 89.8 | 362 | 1 | 1B03_GORGO | | P30381 gorilla gor |
| 5 | 44 | 89.8 | 362 | 1 | 1B15_HUMAN | | P10317 homo sapien |
| 6 | 44 | 89.8 | 362 | 1 | 1B47_HUMAN | | P30487 homo sapien |
| 7 | 44 | 89.8 | 362 | 1 | 1B49_HUMAN | | P18464 homo sapien |
| 8 | 44 | 89.8 | 362 | 1 | 1B52_HUMAN | | P30489 homo sapien |
| 9 | 44 | 89.8 | 362 | 1 | 1B53_HUMAN | | P30490 homo sapien |
| 10 | 44 | 89.8 | 362 | 1 | 1B54_HUMAN | | P30491 homo sapien |
| 11 | 44 | 89.8 | 362 | 1 | 1B60_HUMAN | | P18465 homo sapien |
| 12 | 44 | 89.8 | 362 | 1 | 1B61_HUMAN | | P30497 homo sapien |
| 13 | 44 | 89.8 | 362 | 1 | 1B62_HUMAN | | P10319 homo sapien |
| 14 | 44 | 89.8 | 362 | 1 | 1HLAH_HUMAN | | P01893 homo sapien |
| 15 | 44 | 89.8 | 365 | 1 | 1A23_HUMAN | | P30447 homo sapien |
| 16 | 44 | 89.8 | 365 | 1 | 1A24_HUMAN | | P05534 homo sapien |
| 17 | 39 | 79.6 | 338 | 1 | 1B20_HUMAN | | P30467 homo sapien |
| 18 | 39 | 79.6 | 361 | 1 | 1B14_HUMAN | | P03989 homo sapien |
| 19 | 39 | 79.6 | 362 | 1 | 1B05_HUMAN | | P30461 homo sapien |
| 20 | 39 | 79.6 | 362 | 1 | 1B16_HUMAN | | P19373 homo sapien |
| 21 | 39 | 79.6 | 362 | 1 | 1B18_HUMAN | | P10318 homo sapien |
| 22 | 39 | 79.6 | 362 | 1 | 1B19_HUMAN | | Q08136 homo sapien |
| 23 | 39 | 79.6 | 362 | 1 | 1B29_HUMAN | | P18463 homo sapien |
| 24 | 39 | 79.6 | 362 | 1 | 1B41_HUMAN | | P30481 homo sapien |
| 25 | 39 | 79.6 | 362 | 1 | 1B42_HUMAN | | P30482 homo sapien |
| 26 | 39 | 79.6 | 362 | 1 | 1B45_HUMAN | | P30485 homo sapien |
| 27 | 39 | 79.6 | 365 | 1 | 1A25_HUMAN | | P18462 homo sapien |
| 28 | 39 | 79.6 | 365 | 1 | 1A32_HUMAN | | P10314 homo sapien |
| 29 | 35 | 71.4 | 195 | 1 | TP11_BOVIN | | P15694 bos taurus |
| 30 | 34 | 69.4 | 359 | 1 | 1B40_HUMAN | | P10320 homo sapien |
| 31 | 34 | 69.4 | 365 | 1 | 1A04_GORGO | | P30378 gorilla gor |
| 32 | 34 | 69.4 | 435 | 1 | K1CH_RAT | | Q01134 rattus norv |
| 33 | 34 | 69.4 | 506 | 1 | CP5N_CANTR | | P30611 candida tro |
| 34 | 33 | 67.3 | 244 | 1 | YEHT_ECOLI | | P33356 escherichia |

```

SQ SEQUENCE 359 AA; 40173 MW; 5395FFC9 CRC32;

Query Match      89.8%; Score 44; DB 1; Length 359;
Best Local Similarity 90.0%; Pred. No. 0.077;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRILLRY 10
   ||||| |||
Db 95 RENLRIALRY 104

RESULT 2
1B01_GORGO STANDARD; PRT; 362 AA.
AC P30379;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0101 ALPHA CHAIN PRECURSOR.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92078860.
RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;
RT "Gorilla class I major histocompatibility complex alleles: comparison
to human and chimpanzee class I.";
RL J. Exp. Med. 174:1491-1509(1991).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60255; CAA42807.1; -
CC PIR; JH0539; JH0539.
CC HSSP; P03989; 1HSA.
CC PROSITE; PS00290; IG_MHC; 1.
CC PFAM; PF00047; ig; 1.
CC PFAM; PF00129; MHC_I; 1.
CC MHC I; Transmembrane; Glycoprotein; Signal.
CC SIGNAL 1 24 BY SIMILARITY.
CC CHAIN 25 362 CLASS I HISTOCOMPATIBILITY ANTIGEN,
GOGO-B0101 ALPHA CHAIN.
CC DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
CC DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
CC DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
CC DOMAIN 299 308 CONNECTING PEPTIDE.
CC TRANSMEM 309 332
CC DOMAIN 333 362 CYTOPLASMIC TAIL.
CC DISULFID 125 188 BY SIMILARITY.
CC DISULFID 227 283 BY SIMILARITY.
CC CARBOHYD 110 110 BY SIMILARITY.
SQ SEQUENCE 362 AA; 40170 MW; 2E33E2B8 CRC32;

Query Match      89.8%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.077;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRILLRY 10
   ||||| |||
Db 99 RENLRIALRY 108

RESULT 3
1B02_GORGO STANDARD; PRT; 362 AA.
AC P30380;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0102 ALPHA CHAIN PRECURSOR.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92078860.
RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;
RT "Gorilla class I major histocompatibility complex alleles: comparison
to human and chimpanzee class I.";
RL J. Exp. Med. 174:1491-1509(1991).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60693; CAA43101.1; -
CC PIR; JH0540; JH0540.
CC HSSP; P03989; 1HSA.
CC PROSITE; PS00290; IG_MHC; 1.
CC PFAM; PF00047; ig; 1.
CC PFAM; PF00129; MHC_I; 1.
CC MHC I; Transmembrane; Glycoprotein; Signal.
CC SIGNAL 1 24 BY SIMILARITY.
CC CHAIN 25 362 CLASS I HISTOCOMPATIBILITY ANTIGEN,
GOGO-B0102 ALPHA CHAIN.
CC DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
CC DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
CC DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
CC DOMAIN 299 308 CONNECTING PEPTIDE.
CC TRANSMEM 309 332
CC DOMAIN 333 362 CYTOPLASMIC TAIL.
CC DISULFID 125 188 BY SIMILARITY.
CC DISULFID 227 283 BY SIMILARITY.
CC CARBOHYD 110 110 BY SIMILARITY.
SQ SEQUENCE 362 AA; 40204 MW; 3CF119AD CRC32;

Query Match      89.8%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.077;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLRILLRY 10
   ||||| |||
Db 99 RENLRIALRY 108

RESULT 4
1B03_GORGO STANDARD; PRT; 362 AA.
AC P30381;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE CLASS I HISTOCOMPATIBILITY ANTIGEN, GOGO-B0103 ALPHA CHAIN PRECURSOR.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

```

OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92078860.
 RA LORR D.A., WARREN E., TAYLOR P., PARHAM P.;
 RT "Gorilla class I major histocompatibility complex alleles: comparison
 to human and chimpanzee class I.";
 RL J. Exp. Med. 174:1491-1509(1991).
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 THE IMMUNE SYSTEM.
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 MICROGLOBULIN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X60254; CAA42806.1; -
 DR PIR; JH0541; JH0541.
 DR HSSP; P03989; 1HSA.
 DR PFAM; PF00047; 1g; 1.
 DR MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT CLASS I HISTOCOMPATIBILITY ANTIGEN,
 FT GOGO-B0103 ALPHA CHAIN.
 FT EXTRACELLULAR ALPHA-1.
 FT EXTRACELLULAR ALPHA-2.
 FT EXTRACELLULAR ALPHA-3.
 FT CONNECTING PEPTIDE.
 FT CYTOPLASMIC TAIL.
 FT BY SIMILARITY.
 FT DISULFID 125 188
 FT DISULFID 227 283
 FT CARBOHYD 110 110
 FT BY SIMILARITY.
 SQ SEQUENCE 362 AA; 40248 MW; FEA6A941 CRC32;
 Query Match 89.8%; Score 44; DB 1; Length 362;
 Best Local Similarity 90.0%; Pred. No. 0.077;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 RENTRILLY 10
 Db 99 RENTRILLY 108
 RESULT 5
 ID 1B15_HUMAN STANDARD; PRT; 362 AA.
 AC P10317;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-27 B*2702 ALPHA CHAIN
 DE PRECURSOR (B-27K) (B27.2).
 GN HLA-B OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86220133.
 RA SEEMANN G.H.A., REIN R.S., BROWN C.S., PLOEGH H.L.;
 RT "Gene conversion-like mechanisms may generate polymorphism in human
 class I genes.";
 RL EMBO J. 5:547-552(1986).
 RN [2]

RP SEQUENCE FROM N.A.
 RA PARHAM P., ARNETT K.L., ADAMS E.J.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 86-107 AND 171-181.
 RX MEDLINE; 86042671.
 RA VEGA M.A., EZQUERRA A., ROJO S., APARICIO P., BRAGADO R.,
 RA LOPEZ DE CASTRO J.A.;
 RT "Structural analysis of an HLA-B*27 functional variant: identification
 of residues that contribute to the specificity of recognition by
 cytolytic T lymphocytes.";
 RT Proc. Natl. Acad. Sci. U.S.A. 82:7394-7398(1985).
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 THE IMMUNE SYSTEM.
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 MICROGLOBULIN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X03664; CAA27301.1; -
 DR EMBL; X03667; CAA27301.1; JOINED.
 DR EMBL; L38504; AAA69724.1; -
 DR PIR; B25092; HLHUBK.
 DR HSSP; P03989; 1HSA.
 DR MIN; 142830; -
 DR PROSITE; PS00290; 1g; MHC; 1.
 DR PFAM; PF00047; 1g; 1.
 DR PFAM; PF00129; MHC.I; 1.
 DR MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
 FT B-27 B*2702 ALPHA CHAIN.
 FT EXTRACELLULAR ALPHA-1.
 FT EXTRACELLULAR ALPHA-2.
 FT EXTRACELLULAR ALPHA-3.
 FT CONNECTING PEPTIDE.
 FT CYTOPLASMIC TAIL.
 FT BY SIMILARITY.
 FT DISULFID 125 188
 FT DISULFID 227 283
 FT BY SIMILARITY.
 SQ SEQUENCE 362 AA; 40397 MW; 9798F0BB CRC32;
 Query Match 89.8%; Score 44; DB 1; Length 362;
 Best Local Similarity 90.0%; Pred. No. 0.077;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 RENTRILLY 10
 Db 99 RENTRILLY 108
 RESULT 6
 ID 1B47_HUMAN STANDARD; PRT; 362 AA.
 AC P30487;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-49(B-21) B*4901 ALPHA CHAIN
 DE PRECURSOR.
 GN HLA-B OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]

HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-52(B-5) B*5201 ALPHA CHAIN
 DE PRECURSOR.

DE DE
 GN HLA-B OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE; 92269955.
 RA HAYASHI H., ENNIS P.D., ARIGA H., SALTER R.D., PARHAM P., KANO K.,
 RA TAKIGUCHI M.;
 RT "HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the
 RT helical region of the alpha 1 domain.";
 RL J. Immunol. 142:306-311(1989).
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M22799; AAA59645.1; ALT SEQ.
 DR EMBL; M22793; AAA59645.1; JOINED.
 DR EMBL; M22794; AAA59645.1; JOINED.
 DR EMBL; M22795; AAA59645.1; JOINED.
 DR EMBL; M22796; AAA59645.1; JOINED.
 DR EMBL; M22797; AAA59645.1; JOINED.
 DR EMBL; M22798; AAA59645.1; JOINED.
 DR PIR; B30345; B30345.
 DR PIR; B30548; B30548.
 DR HSP; P30491; IALM.
 DR MIM; 142830; .
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; Ig; 1.
 DR PFAM; PF00129; MHC.I; 1.
 KW MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
 FT B-51(B-5) B*5104 ALPHA CHAIN.
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT DOMAIN 309 332
 FT TRANSMEM 333 362
 FT DOMAIN 333 362
 FT CARBOHYD 110 110
 FT BY SIMILARITY.
 FT DISULFID 125 188
 FT BY SIMILARITY.
 FT DISULFID 227 283
 FT BY SIMILARITY.
 SQ SEQUENCE 362 AA; 40560 MW; F22F08AB CRC32;

Query Match 89.8%; Score 44; DB 1; Length 362;
 Best Local Similarity 90.0%; Pred. No. 0.077;
 Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

OY 1 RENRILLRY 10
 Db 99 RENRILLRY 108
 RESULT 9
 ID 1B53_HUMAN STANDARD; PRT; 362 AA.
 AC P30490;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)

DE DE
 GN HLA-B OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE; 89080265.
 RA HAYASHI H., ENNIS P.D., ARIGA H., SALTER R.D., PARHAM P., KANO K.,
 RA TAKIGUCHI M.;
 RT "HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the
 RT helical region of the alpha 1 domain.";
 RL J. Immunol. 142:306-311(1989).
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M22799; AAA59645.1; ALT SEQ.
 DR EMBL; M22793; AAA59645.1; JOINED.
 DR EMBL; M22794; AAA59645.1; JOINED.
 DR EMBL; M22795; AAA59645.1; JOINED.
 DR EMBL; M22796; AAA59645.1; JOINED.
 DR EMBL; M22797; AAA59645.1; JOINED.
 DR EMBL; M22798; AAA59645.1; JOINED.
 DR PIR; B30345; B30345.
 DR PIR; B30548; B30548.
 DR HSP; P30491; IALM.
 DR MIM; 142830; .
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; Ig; 1.
 DR PFAM; PF00129; MHC.I; 1.
 KW MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
 FT B-52(B-5) B*5201 ALPHA CHAIN.
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT DOMAIN 309 332
 FT TRANSMEM 333 362
 FT DOMAIN 333 362
 FT CARBOHYD 110 110
 FT BY SIMILARITY.
 FT DISULFID 125 188
 FT BY SIMILARITY.
 FT DISULFID 227 283
 FT BY SIMILARITY.
 SQ SEQUENCE 362 AA; 40521 MW; 3B436FE8 CRC32;

Query Match 89.8%; Score 44; DB 1; Length 362;
 Best Local Similarity 90.0%; Pred. No. 0.077;
 Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

OY 1 RENRILLRY 10
 Db 99 RENRILLRY 108
 RESULT 10
 ID 1B54_HUMAN STANDARD; PRT; 362 AA.
 AC P30491;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-53 B*5301 ALPHA CHAIN
DE PRECURSOR.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91033941.
RA HAYASHI H., Ooba T., NAKAYAMA S., SEKIMATA M., KANO K.,
RA TAKIGUCHI M.;
RT "Allostericities between HLA-B*53 and HLA-B*35 are generated by
RT substitution of the residues associated with HLA-B*4/B*6 public
RT epitopes.";
RL Immunogenetics 32:195-199(1990).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 25-302.
RX MEDLINE; 96209672.
RA SMITH K.J., REID S.W., HARLOS K., MCMICHAEL A.J., STUART D.I.,
RA BELL J.I., JONES E.Y.;
RT "Bound water structure and polymorphic amino acids act together to
RT allow the binding of different peptides to MHC class I HLA-B*53.";
RL Immunity 4:215-228(1996).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M58636; AAA36228.1; -;
DR PIR; A45834; A45834.
DR PDB; 1A10; 08-APR-98.
DR PDB; 1A10; 08-APR-98.
DR MIM; 142830; -;
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_1; 1.
DR MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT BW-53 B*5301 ALPHA CHAIN.
FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
FT DOMAIN 299 308 CONNECTING PEPTIDE.
FT TRANSMEM 309 332
FT DOMAIN 333 362 CYTOPLASMIC TAIL.
FT CARBOHYD 110 110 BY SIMILARITY.
FT DISULFID 125 188
FT DISULFID 227 283
SQ SEQUENCE 362 AA; 40495 MW; 2BDC746E CRC32;

Query Match 89.8%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.077;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RENLIRLLY 10
Db 99 RENLIRLLY 108

RESULT 11
ID 1B60_HUMAN STANDARD; PRT; 362 AA.
AC P18465.

DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B*57(B-17) B*5701 ALPHA
DE CHAIN PRECURSOR (BW57.1).
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90207291.
RA ENNIS P.D., ZEMOUR J., SALTER R.D., PARHAM P.;
RT "Rapid cloning of HLA-A,B cDNA by using the polymerase chain
RT reaction: frequency and nature of errors produced in amplification.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91067476.
RA ISAMAT M., GIRDLESTONE J., MILSTEIN C.;
RT "Nucleotide sequence of an HLA-B*57 gene.";
RL Nucleic Acids Res. 18:6702-6702(1990).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M32318; AAA36231.1; -;
DR PIR; X55711; CAA39244.1; -;
DR PIR; S12622; S12622.
DR PIR; D35997; D35997.
DR HSSP; P30491; 1ALM.
DR MIM; 142830; -;
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_1; 1.
DR MHC I; Transmembrane; Glycoprotein; Signal.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT B*57(B-17) B*5701 ALPHA CHAIN.
FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
FT DOMAIN 299 308 CONNECTING PEPTIDE.
FT TRANSMEM 309 332
FT DOMAIN 333 362 CYTOPLASMIC TAIL.
FT CARBOHYD 110 110 BY SIMILARITY.
FT DISULFID 125 188 BY SIMILARITY.
FT DISULFID 227 283 BY SIMILARITY.
SQ SEQUENCE 362 AA; 40224 MW; D91DF8DD CRC32;

Query Match 89.8%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.077;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RENLIRLLY 10
Db 99 RENLIRLLY 108

RESULT 12
ID 1B61_HUMAN STANDARD; PRT; 362 AA.
AC P30497;

DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-57(B-17) B*5702 ALPHA CHAIN
 DE PRECURSOR (BW57.2).
 GN HLA-B OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 93056508.
 RA MADRIGAL J.A., SELITCH M.P., HILDEBRAND W.H., BENJAMIN R.J.,
 RA LITTLE A.M., ZEMMOUR J., ENNIS P.D., WARD F.E., PETZL-ERLER M.L.,
 RA MARTEL R.W., DU TOIT E.D., PARHAM P.;
 RT "Distinctive HLA-A,B antigens of black populations formed by
 RT interallelic conversion."
 RL J. Immunol. 149:3411-3415(1992).
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC THE IMMUNE SYSTEM.
 CC MICROGLOBULIN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X61707; CAA43876.1; -
 DR PIR; S16774; S16774.
 DR HSP; P30491; IALM.
 DR MIM; 142830; -
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; Ig; 1.
 DR MHC I; Transmembrane; Glycoprotein; Signal.
 KW MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT TRANSMEM 309 332
 FT DOMAIN 333 362
 FT CARBOHYD 110 110
 FT DISULFID 125 188
 FT DISULFID 227 283
 SQ SEQUENCE 362 AA; 40342 MW; 628C2156 CRC32;

Query Match 89.8%; Score 44; DB 1; Length 362;
 Best Local Similarity 90.0%; Pred. No. 0.077;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLIRLLY 10
 Db 99 RENLIRLLY 108

RESULT 13
 ID 1862_HUMAN STANDARD; PRT; 362 AA.
 AC P10319;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-58(B-17) B*5801 ALPHA
 DE CHAIN PRECURSOR.
 GN HLA-B OR HLAB.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86008247.
 RA WAYS J.P., COPPIN H.L., PARHAM P.;
 RT "The complete primary structure of HLA-Bw58."
 RT J. Biol. Chem. 260:11924-11933(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD;
 RA INOUE T., OGAWA A.;
 RA Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M11799; AAA59628.1; -
 DR EMBL; AB008102; BAA22916.1; -
 DR PIR; A23895; HLHUB8.
 DR HSP; P30491; IALM.
 DR MIM; 142830; -
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; Ig; 1.
 DR MHC I; Transmembrane; Glycoprotein; Signal.
 KW MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT TRANSMEM 309 332
 FT DOMAIN 333 362
 FT CARBOHYD 110 110
 FT DISULFID 125 188
 FT DISULFID 227 283
 SQ SEQUENCE 362 AA; 40337 MW; 3E5E7534 CRC32;

Query Match 89.8%; Score 44; DB 1; Length 362;
 Best Local Similarity 90.0%; Pred. No. 0.077;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLIRLLY 10
 Db 99 RENLIRLLY 108

RESULT 14
 ID HLAH_HUMAN STANDARD; PRT; 362 AA.
 AC P01893;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 13-DEC-1999 (Rel. 39, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, ALPHA CHAIN H PRECURSOR
 DE (HLA-AR) (HLA-12.4).
 GN HLA-H OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

RN  SEQUENCE FROM N.A.
RX  MEDLINE: 82151002.
RA  MALISSEN M., MALISSEN B., JORDAN B.R.;
RT  "Exon/intron organization and complete nucleotide sequence of an HLA
RT  gene."
RL  Proc. Natl. Acad. Sci. U.S.A. 79:893-897(1982).
CC  -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC  THE IMMUNE SYSTEM. COULD BE THE PRODUCT OF A PSEUDOGENE.
CC  -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC  MICROGLOBULIN).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: J00191; AAA36218.1; ALT_INIT.
CC  PIR: A02189; HLH12.
CC  HSSP: P03989; LHS.
CC  MIM: 142800; -.
CC  PROSITE: PS00290; IG_MHC; 1.
CC  PFAM: PF00047; Ig; 1.
CC  PFAM: PF00129; MHC_I; 1.
CC  MHC I; Transmembrane; Glycoprotein; Signal.
CC  SIGNAL 1 24
CC  CHAIN 25 362
CC  HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
CC  ALPHA CHAIN H.
CC  DOMAIN 25 114
CC  EXTRACELLULAR ALPHA-1.
CC  DOMAIN 115 206
CC  EXTRACELLULAR ALPHA-2.
CC  DOMAIN 207 298
CC  EXTRACELLULAR ALPHA-3.
CC  DOMAIN 299 308
CC  CONNECTING PEPTIDE.
CC  TRANSMEM 309 332
CC  CYTOPLASMIC TAIL.
CC  DOMAIN 333 362
CC  CARBOHYD 110 110
CC  BY SIMILARITY.
CC  DISULFID 227 283
CC  BY SIMILARITY.
CC  SEQUENCE 362 AA; 40850 MW; 5E610F63 CRC32;
SQ

```

```

Query Match      89.8%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.077;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy  1 RNLRLRLRY 10
    ||||| |||
Db  99 RNLRLRLRY 108

```

```

RESULT 15
1A23_HUMAN
ID  1A23_HUMAN STANDARD; PRT; 365 AA.
AC  P30447;
DT  01-APR-1993 (Rel. 25, Created)
DT  01-APR-1993 (Rel. 25, Last sequence update)
DT  15-JUL-1999 (Rel. 38, Last annotation update)
DE  HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-23(A-9) ALPHA CHAIN
DE  PRECURSOR.
GN  HLA-A OR HLA.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC  Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN  [1]
RP  SEQUENCE FROM N.A. (A*2301).
RX  MEDLINE: 92104637.
RA  LITTLE A.-M., MADRIGAL J.A., PARHAM P.;
RT  "Molecular definition of an elusive third HLA-A9 molecule: HLA-A9.3."
RL  Immunogenetics 35:41-45(1992).
CC  -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC  THE IMMUNE SYSTEM.
CC  -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-

```

```

CC  MICROGLOBULIN).
CC  -!- POLYMORPHISM: THE ONLY ALLELE OF A-23 KNOWN IS A*2301 WHICH IS
CC  SHOWN HERE.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: M64742; AAA03662.1; -.
CC  HSSP: P01892; IAQD.
CC  MIM: 142800; -.
CC  PROSITE: PS00290; IG_MHC; 1.
CC  PFAM: PF00047; Ig; 1.
CC  PFAM: PF00129; MHC_I; 1.
CC  MHC I; Transmembrane; Glycoprotein; Signal.
CC  SIGNAL 1 24
CC  CHAIN 25 365
CC  HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
CC  A-23(A-9) ALPHA CHAIN.
CC  EXTRACELLULAR ALPHA-1.
CC  DOMAIN 115 206
CC  EXTRACELLULAR ALPHA-2.
CC  DOMAIN 207 298
CC  EXTRACELLULAR ALPHA-3.
CC  DOMAIN 299 308
CC  CONNECTING PEPTIDE.
CC  TRANSMEM 309 332
CC  CYTOPLASMIC TAIL.
CC  DOMAIN 333 365
CC  CARBOHYD 110 110
CC  BY SIMILARITY.
CC  DISULFID 125 188
CC  BY SIMILARITY.
CC  DISULFID 227 283
CC  BY SIMILARITY.
CC  SEQUENCE 365 AA; 40732 MW; B1C21094 CRC32;
SQ

```

```

Query Match      89.8%; Score 44; DB 1; Length 365;
Best Local Similarity 90.0%; Pred. No. 0.078;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy  1 RNLRLRLRY 10
    ||||| |||
Db  99 RNLRLRLRY 108

```

```

Search completed: February 8, 2000, 00:59:51
Job time: 3780 sec

```


SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 250)
AUTHORS Rojas-Munoz,A., Mendez,I. and Yunis,I.
TITLE Molecular evolution of HLA-B locus in a small population amerindian community :The Nukak-Maku
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 250)
AUTHORS Rojas-Munoz,A.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1996) A. Rojas-Munoz, National Institute Of Health, Immunogenetics, Avenida El Dorado, Carrera 50, Santafe De Bogota / Zona 6, COLOMBIA
FEATURES Location/Qualifiers
source 1..250
/organism="Homo sapiens"
/isolate="Isabel-26"
/isolate="from amerindian community Nukak-Maku"
/db_xref="taxon:9606"
/chromosome="6"
/dev_stage="adult"
/tissue_type="blood"
/cell_type="white"
/lab_host="E.coli TGI"
/clone="CHBC2(b)"
/clone="CHBC3(b)"
/clone="CHBC4(b)"
14..250
/gene="HLA-B"
<14..>250
/gene="HLA-B"
/note="allele HB(b)"
/number=2
BASE COUNT 51 a 78 c 87 g 34 t
ORIGIN
alignment_scores:
Quality: 44.00 Length: 10
Ratio: 4.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 90.000
alignment_block:
US-08-653-294-10 x HSHLABHBB ..
Align seg 1/1 to: HSHLABHBB from: 1 to: 250
1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
|||||
209 CGAGAGAACTCGGATCGGCTCGCTAC 238
seq_name: gb_prl: HSHLABHBB
seq_documentation_block:
LOCUS HSHLABHBB 250 bp DNA PRI 10-OCT-1996
DEFINITION H. sapiens HLA-B gene, exon 2, HB(d) allele.
ACCESSION Y08694
VERSION Y08694.1 GI:1619289
KEYWORDS HLA-B gene; human leukocyte antigen.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 250)
AUTHORS Rojas-Munoz,A., Mendez,I. and Yunis,I.
TITLE Molecular evolution of HLA-B locus in a small population amerindian community :The Nukak-Maku
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 250)
AUTHORS Rojas-Munoz,A.
TITLE Direct Submission

JOURNAL Submitted (07-OCT-1996) A. Rojas-Munoz, National Institute Of Health, Immunogenetics, Avenida El Dorado, Carrera 50, Santafe De Bogota / Zona 6, COLOMBIA
FEATURES Location/Qualifiers
source 1..250
/organism="Homo sapiens"
/isolate="Norman-51"
/isolate="from amerindian community Nukak-Maku"
/db_xref="taxon:9606"
/chromosome="6"
/dev_stage="adult"
/tissue_type="blood"
/cell_type="white"
/lab_host="E.coli TGI"
/clone="CHBC1(d)"
14..250
/gene="HLA-B"
<14..>250
/gene="HLA-B"
/note="allele HB(d)"
/number=2
BASE COUNT 58 a 78 c 79 g 35 t
ORIGIN
alignment_scores:
Quality: 44.00 Length: 10
Ratio: 4.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 90.000
alignment_block:
US-08-653-294-10 x HSHLABHBD ..
Align seg 1/1 to: HSHLABHBD from: 1 to: 250
1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
|||||
209 CGAGAGAACTCGGATCGGCTCGCTAC 238
seq_name: gb_pr2: HSHLABB1
seq_documentation_block:
LOCUS HSHLABB1 250 bp DNA PRI 22-MAR-1997
DEFINITION Human cell line THAI DCH010 MHC class I HLA-B gene (allele HLA-B*1513), exon 2.
ACCESSION U90420
VERSION U90420.1 GI:1905830
KEYWORDS
SEGMENT 1 of 2
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 250)
AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D., Rungroung,E., Bejchandra,S., Juji,T. and Tokunaga,K.
TITLE B15 alleles (B*1513)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 250)
AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D., Rungroung,E., Bejchandra,S., Juji,T. and Tokunaga,K.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine, Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700, Thailand
FEATURES Location/Qualifiers
source 1..250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="VI"
/map="6p21"
/cell_type="lymphoblastoid"
/cell_line="THAI DCH010"


```

VERSION      U90615.1  GI:1906037
KEYWORDS
SEGMENT      1 of 2
SOURCE       human.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
              Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE     1 (bases 1 to 250)
AUTHORS      Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
              Siriboonrit,U., Rungroung,E. and Bejchandra,S.
TITLE        B*51V alleles
JOURNAL      Unpublished
REFERENCE     2 (bases 1 to 250)
AUTHORS      Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
              Siriboonrit,U., Rungroung,E. and Bejchandra,S.
TITLE        Direct Submission
JOURNAL      Submitted (25-FEB-1997) Transfusion Medicine, Faculty of Medicine,
              Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
              Thailand
FEATURES      Location/Qualifiers
              source
                1..250
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="6"
                /map="6p21"
                /cell_type="lymphoblastoid"
                /cell_line="THAI DCH011"
                /gene="HLA-B"
                /note="Allele: HLA-B*513"
                /number=2
                /product="MHC class I HLA-B"
              exon
                55 a      83 c      80 g      32 t
              BASE COUNT
              ORIGIN
alignment_scores:
  Quality: 44.00      Length: 10
  Ratio: 4.889      Gaps: 0
  Percent Similarity: 90.000      Percent Identity: 90.000
alignment_block:
  US-08-653-294-10 x HSHLAB11
  Align seg 1/1 to: HSHLAB11 from: 1 to: 250
  1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
  |||||
  202 CGAGAGAACTCGGATCGGCTCGCTAC 231
seq_name: gb_pr2:HSHLABJ1

seq_documentation_block:
LOCUS      HSHLABJ1      250 bp      DNA      PRI      22-MAR-1997
DEFINITION Human cell line THAI DCH028 MHC class I HLA-B gene (allele
              HLA-B*1513), exon 2.
ACCESSION  U90424
VERSION     U90424.1  GI:1905838
KEYWORDS    1 of 2
SEGMENT     human.
SOURCE      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
              Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 250)
AUTHORS     Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
              Rungroung,E. and Bejchandra,S.
TITLE       B15 alleles (B*1513)
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 250)
AUTHORS     Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
              Rungroung,E. and Bejchandra,S.
TITLE       Direct Submission
JOURNAL     Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine,
              Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
              Thailand
FEATURES      Location/Qualifiers
              source
                1..250
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="VI"
                /map="6p21"
                /cell_type="lymphoblastoid"
                /cell_line="THAI DCH028"
                /gene="HLA-B"
                /note="Allele: HLA-B*1513"
                /number=2
              exon
                55 a      82 c      80 g      32 t
              BASE COUNT
              ORIGIN
alignment_scores:
  Quality: 44.00      Length: 10
  Ratio: 4.889      Gaps: 0
  Percent Similarity: 90.000      Percent Identity: 90.000
alignment_block:
  US-08-653-294-10 x HSHLABG1
  Align seg 1/1 to: HSHLABG1 from: 1 to: 250
  1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
  |||||
  202 CGAGAGAACTCGGATCGGCTCGCTAC 231
seq_name: gb_pr2:HSHLABI1

seq_documentation_block:
LOCUS      HSHLABI1      250 bp      DNA      PRI      22-MAR-1997
DEFINITION Human cell line THAI DCH011 MHC class I HLA-B gene (allele
              HLA-B*1513), exon 2.
ACCESSION  U90422
VERSION     U90422.1  GI:1905834
KEYWORDS    1 of 2
SEGMENT     human.
SOURCE      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
              Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 250)
AUTHORS     Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
              Rungroung,E., Bejchandra,S., Blasczyk,R. and Grosse-Wilde,H.
TITLE       B15 alleles (B*1513)
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 250)
AUTHORS     Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
              Rungroung,E., Bejchandra,S., Blasczyk,R. and Grosse-Wilde,H.

```


BASE COUNT 55 a 83 c 80 g 32 t
ORIGIN

alignment_scores:
Quality: 44.00 Length: 10
Ratio: 4.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:
US-08-653-294-10 x HSHLABT1 ..

Align seg 1/1 to: HSHLABT1 from: 1 to: 250

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
|||||
202 CGAGAGAACCTGCGGATCGGCTCCGCTAC 231

seq_name: gb_pr2:HSHLABT1

seq_documentation_block: 250 bp DNA PRI 22-MAR-1997
LOCUS HSHLABT1
DEFINITION Human cell line THAI DCH009 MHC class I HLA-B gene (allele
HLA-B*1513), exon 2.

ACCESSION U90418
VERSION U90418.1 GI:1905826

KEYWORDS
SEGMENT
SOURCE

1 of 2

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 250)

AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,

Rungroung,E., Bejchandra,S., Juji,T. and Tokunaga,K.

TITLE B15 alleles (B*1513)

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 250)

AUTHORS Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,

Rungroung,E., Bejchandra,S., Juji,T. and Tokunaga,K.

TITLE Direct Submission

JOURNAL Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine,

Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,

Thailand

FEATURES

Location/Qualifiers

1..250

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="VI"

/map="6p21"

/cell_type="lymphoblastoid"

/cell_line="THAI DCH009"

1..250

/gene="HLA-B"

/note="Allele: HLA-B*1513"

/number=2

/product="MHC class I HLA-B"

BASE COUNT 55 a 83 c 80 g 32 t

ORIGIN

alignment_scores:
Quality: 44.00 Length: 10
Ratio: 4.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:
US-08-653-294-10 x HSHLABT1 ..

Align seg 1/1 to: HSHLABT1 from: 1 to: 250

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10

|||||
202 CGAGAGAACCTGCGGATCGGCTCCGCTAC 231

seq_name: gb_pr4:AF022159

seq_documentation_block: 255 bp DNA PRI 05-JAN-1999
LOCUS AF022159
DEFINITION Homo sapiens isolate 026 MHC class I antigen HLA-H (HLA-H)
pseudogene, partial sequence.

ACCESSION AF022159
VERSION AF022159.1 GI:2655062

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 255)

AUTHORS Grimsley,C., Mather,K.A. and Ober,C.

TITLE HLA-H: a pseudogene with increased variation due to balancing

selection at neighboring loci

JOURNAL Mol. Biol. Evol. 15 (12), 1581-1588 (1998)

MEDLINE 99083426

REFERENCE 2 (bases 1 to 255)

AUTHORS Grimsley,C., Mather,K.A. and Ober,C.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-1997) Fred Hutchinson Cancer Research Center,

1100 Fairview Ave. N., M374, Seattle, WA 98109, USA

FEATURES

Location/Qualifiers

1..255

/organism="Homo sapiens"

/isolate="026"

/db_xref="taxon:9606"

/chromosome="6"

/map="6p21.3"

/note="African-American individual"

1..255

/gene="HLA-H"

/pseudo

/number=2

<1..>255

/gene="HLA-H"

/codon_start=1

/pseudo

<1..>255

/gene="HLA-H"

/pseudo

BASE COUNT 50 a 81 c 86 g 38 t

ORIGIN

alignment_scores:
Quality: 44.00 Length: 10
Ratio: 4.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:
US-08-653-294-10 x AF022159 ..

Align seg 1/1 to: AF022159 from: 1 to: 255

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10

|||||
223 CGAGAGAACCTGCGGATCGGCTCCGCTAC 252

seq_name: gb_pr4:AF022160

seq_documentation_block:
LOCUS AF022160 259 bp DNA PRI 05-JAN-1999
DEFINITION Homo sapiens isolate 034 MHC class I antigen HLA-H (HLA-H)
pseudogene, partial sequence.

ACCESSION AF022160
VERSION AF022160.1 GI:2655063

```

KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 259)
AUTHORS     Grimsley,C., Mather,K.A. and Ober,C.
TITLE       HLA-H: a pseudogene with increased variation due to balancing
            selection at neighboring loci
JOURNAL     Mol. Biol. Evol. 15 (12), 1581-1588 (1998)
MEDLINE     99083426
REFERENCE   2 (bases 1 to 259)
AUTHORS     Grimsley,C., Mather,K.A. and Ober,C.
TITLE       Direct Submission
JOURNAL     Submitted (03-SEP-1997) Fred Hutchinson Cancer Research Center,
            1100 Fairview Ave. N., M374, Seattle, WA 98109, USA
FEATURES
            Location/Qualifiers
            source
            1..259
            /organism="Homo sapiens"
            /isolate="034"
            /db_xref="taxon:9606"
            /chromosome="6"
            /map="6p21.3"
            /note="African-American individual"
            1..259
            /gene="HLA-H"
            /pseudo
            /number=2
            <1..>259
            /gene="HLA-H"
            /note="MHC class I antigen HLA-H"
            /codon_start=1
            /pseudo
            <1..>259
            /gene="HLA-H"
            /pseudo
BASE COUNT  52 a 83 c 85 g 37 t 2 others
ORIGIN

alignment_scores:
    Quality: 44.00      Length: 10
    Ratio: 4.889       Gaps: 0
    Percent Similarity: 90.000      Percent Identity: 90.000

alignment_block:
US-08-653-294-10 x AF022160 ..
Align seg 1/1 to: AF022160 from: 1 to: 259
1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
|||||
223 CGAGAGAACCTGCGGATCGCGCTCGCTAC 252

seq_name: gb_pat:114590

seq_documentation_block:
LOCUS      114590          270 bp      DNA
DEFINITION Sequence 67 from patent US 5451512.
ACCESSION  114590
VERSION    114590.1 GI:997073
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 270)
AUTHORS   Apple,R.J., Bugawan,T.L. and Erlich,H.A.
TITLE     Methods and reagents for HLA class I A locus DNA typing
JOURNAL   Patent: US 5451512-A 67 19-SEP-1995;
FEATURES   Location/Qualifiers
            source
            1..270
            /organism="unknown"
BASE COUNT  55 a 84 c 95 g 36 t
ORIGIN

alignment_scores:
    Quality: 44.00      Length: 10
    Ratio: 4.889       Gaps: 0
    Percent Similarity: 90.000      Percent Identity: 90.000

alignment_block:
US-08-653-294-10 x I14591 ..
Align seg 1/1 to: I14591 from: 1 to: 270
1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
|||||
222 CGAGAGAACCTGCGGATCGCGCTCGCTAC 251

seq_name: gb_pat:114592

seq_documentation_block:
LOCUS      114592          270 bp      DNA
DEFINITION Sequence 69 from patent US 5451512.
ACCESSION  114592
VERSION    114592.1 GI:997075
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 270)
AUTHORS   Apple,R.J., Bugawan,T.L. and Erlich,H.A.
TITLE     Methods and reagents for HLA class I A locus DNA typing
JOURNAL   Patent: US 5451512-A 69 19-SEP-1995;
FEATURES   Location/Qualifiers
            source
            1..270
            /organism="unknown"
BASE COUNT  55 a 84 c 95 g 36 t
ORIGIN

```

```

alignment_scores:
  Quality: 44.00      Length: 10
  Ratio: 4.889        Gaps: 0
  Percent Similarity: 90.000  Percent Identity: 90.000

alignment_block:
  US-08-653-294-10 x I14592  ..
  Align seg 1/1  to: I14592  from: 1  to: 270
      1  ArgGluAsnLeuArgIleLeuLeuArgTyr 10
        |||||||||||||||
      222 CGAGAGAACCTGGGATCGCGCTCGCTAC 251
  
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 13:17:37 ; Search time 209.03 Seconds
(without alignments)
3.317 Million cell updates/sec

Title: US-08-653-294-10
Perfect score: 49
Sequence: 1 RENLIRLLRY 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

SPREMBL.12.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 44 | 89.8 | 89 | 7 | 019569 homo sapien |
| 2 | 44 | 89.8 | 90 | 7 | 046697 gorilla gor |
| 3 | 44 | 89.8 | 133 | 7 | 019189 homo sapien |
| 4 | 44 | 89.8 | 137 | 7 | 095533 pan troglod |
| 5 | 44 | 89.8 | 138 | 7 | 078209 homo sapien |
| 6 | 44 | 89.8 | 172 | 7 | 019770 homo sapien |
| 7 | 44 | 89.8 | 172 | 7 | 019774 homo sapien |
| 8 | 44 | 89.8 | 172 | 7 | 019775 homo sapien |
| 9 | 44 | 89.8 | 172 | 7 | 019780 homo sapien |
| 10 | 44 | 89.8 | 172 | 7 | 095364 homo sapien |
| 11 | 44 | 89.8 | 172 | 7 | 019771 homo sapien |
| 12 | 44 | 89.8 | 172 | 7 | 019772 homo sapien |
| 13 | 44 | 89.8 | 172 | 7 | 019773 homo sapien |
| 14 | 44 | 89.8 | 175 | 7 | 029694 homo sapien |
| 15 | 44 | 89.8 | 180 | 7 | 019607 homo sapien |
| 16 | 44 | 89.8 | 180 | 7 | 019608 homo sapien |
| 17 | 44 | 89.8 | 180 | 7 | 019609 homo sapien |
| 18 | 44 | 89.8 | 180 | 7 | 019610 homo sapien |
| 19 | 44 | 89.8 | 180 | 7 | 019611 homo sapien |
| 20 | 44 | 89.8 | 180 | 7 | 019612 homo sapien |

| | | | | | | |
|----|----|------|-----|---|--------|--------------------|
| 21 | 44 | 89.8 | 180 | 7 | 019613 | 019613 homo sapien |
| 22 | 44 | 89.8 | 181 | 7 | 046703 | 046703 homo sapien |
| 23 | 44 | 89.8 | 181 | 7 | 062917 | 062917 homo sapien |
| 24 | 44 | 89.8 | 181 | 7 | 062892 | 062892 homo sapien |
| 25 | 44 | 89.8 | 181 | 7 | 062899 | 062899 homo sapien |
| 26 | 44 | 89.8 | 181 | 7 | 062920 | 062920 homo sapien |
| 27 | 44 | 89.8 | 181 | 7 | 062922 | 062922 homo sapien |
| 28 | 44 | 89.8 | 181 | 7 | 062923 | 062923 homo sapien |
| 29 | 44 | 89.8 | 181 | 7 | 019623 | 019623 homo sapien |
| 30 | 44 | 89.8 | 181 | 7 | 019747 | 019747 homo sapien |
| 31 | 44 | 89.8 | 181 | 7 | 029667 | 029667 homo sapien |
| 32 | 44 | 89.8 | 181 | 7 | 030198 | 030198 homo sapien |
| 33 | 44 | 89.8 | 181 | 7 | 029708 | 029708 homo sapien |
| 34 | 44 | 89.8 | 181 | 7 | 019631 | 019631 homo sapien |
| 35 | 44 | 89.8 | 181 | 7 | 019769 | 019769 homo sapien |
| 36 | 44 | 89.8 | 181 | 7 | 029724 | 029724 homo sapien |
| 37 | 44 | 89.8 | 181 | 7 | 029910 | 029910 homo sapien |
| 38 | 44 | 89.8 | 181 | 7 | 079559 | 079559 homo sapien |
| 39 | 44 | 89.8 | 181 | 7 | 029679 | 029679 homo sapien |
| 40 | 44 | 89.8 | 181 | 7 | 019521 | 019521 homo sapien |
| 41 | 44 | 89.8 | 181 | 7 | 019597 | 019597 homo sapien |
| 42 | 44 | 89.8 | 181 | 7 | 029909 | 029909 homo sapien |
| 43 | 44 | 89.8 | 181 | 7 | 029701 | 029701 homo sapien |
| 44 | 44 | 89.8 | 181 | 7 | 029841 | 029841 homo sapien |
| 45 | 44 | 89.8 | 181 | 7 | 019354 | 019354 gorilla gor |

ALIGNMENTS

RESULT 1
ID 019569 PRELIMINARY; PRT; 89 AA.
AC 019569;
DT 01-JAN-1998 (TREMREL. 05, Created)
DT 01-MAY-1999 (TREMREL. 10, Last sequence update)
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)
DE MHC CLASS I ANTIGEN (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CAO K., BURDETT L., ZHANG G., FERNANDEZ-VINA M.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF017320; AAS70286.2; -
KW MHC.
FT NON_TER 1 1
FT NON_TER 89 89
SQ SEQUENCE 89 AA; 10606 MW; 99D11089 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 89;
Best Local Similarity 90.0%; Pred. No. 0.12;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENLIRLLRY 10
| | | | | | | |
DB 74 RENLIRLLRY 83

RESULT 2
ID 046697 PRELIMINARY; PRT; 90 AA.
AC 046697;
DT 01-JUN-1998 (TREMREL. 06, Created)
DT 01-JUN-1998 (TREMREL. 06, Last sequence update)
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)
DE MHC CLASS I ANTIGEN HLA-H ORTHOLOG (FRAGMENT).
GN HLA-H.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=SHAMBA;
 RA GRIMSLEY C., MATHER K.A., OBER C.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF022172; AAC99794.1; -;
 DR PFAM: PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 90 90
 SQ SEQUENCE 90 AA; 10689 MW; 5E5F2495 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 90;
 Best Local Similarity 90.0%; Pred. No. 0.12;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RNLRLRLRY 10
 ||||| |||
 Db 75 RNLRLRLRY 84

RESULT 3
 OI9189 PRELIMINARY; PRT: 133 AA.
 AC OI9189;
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)
 DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
 DE 01-NOV-1999 (T-EMBLrel. 12, Last annotation update)
 DE MHC CLASS I HISTOCOMPATIBILITY ANTIGEN-B (HLA-B-27KSH) (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEUKOCYTE;
 RA PETERSDORF E.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U18659; AAB60357.1; -;
 DR MIM: 142830; -;
 DR PFAM: PF00129; MHC_I; 1.
 KW MHC I.
 FT NON_TER 1 1
 FT NON_TER 133 133
 SQ SEQUENCE 133 AA; 15491 MW; 3A3BC802 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 133;
 Best Local Similarity 90.0%; Pred. No. 0.18;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RNLRLRLRY 10
 ||||| |||
 Db 27 RNLRLRLRY 36

RESULT 4
 Q95533 PRELIMINARY; PRT: 137 AA.
 AC Q95533;
 DT 01-FEB-1997 (T-EMBLrel. 02, Created)
 DT 01-FEB-1997 (T-EMBLrel. 02, Last sequence update)
 DT 01-NOV-1998 (T-EMBLrel. 08, Last annotation update)
 DE CLASS I HISTOCOMPATIBILITY ANTIGEN (FRAGMENT).
 GN HLA-B.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Pan.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MODRA;

RX MEDLINE: 94285544.
 RA MCADAM S.N., BOYSON J.E., LIU X., GARBER T.L., HUGHES A.L.,
 RA BONTROP R.E., WATKINS D.I.;
 RT "A uniquely high level of recombination at the HLA-B locus."
 RL Natl. Acad. Sci. U.S.A. 91:5893-5897(1994).
 DR EMBL: U05585; AAA50188.1; -;
 DR PFAM: PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 137 137
 SQ SEQUENCE 137 AA; 15922 MW; B316D3BC CRC32;

Query Match 89.8%; Score 44; DB 7; Length 137;
 Best Local Similarity 90.0%; Pred. No. 0.19;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RNLRLRLRY 10
 ||||| |||
 Db 40 RNLRLRLRY 49

RESULT 5
 O78209 PRELIMINARY; PRT: 138 AA.
 AC O78209;
 DT 01-NOV-1998 (T-EMBLrel. 08, Created)
 DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
 DT 01-MAY-1999 (T-EMBLrel. 10, Last annotation update)
 DE HUMAN LEUKOCYTE ANTIGEN PRECURSOR (FRAGMENT).
 GN HLA-A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 98007772.
 RA LAFORET M., FROELICH N., PARISSIADIS A., BAUSINGER H., PFEIFFER B.,
 RA TONGIO M.M.;
 RT "An intronic mutation responsible for a low level of expression of an
 RT HLA-A*24 allele."
 RL Tissue Antigens 50:340-346(1997).
 DR EMBL: 272423; CAA96533.1; -;
 DR PFAM: PF00129; MHC_I; 1.
 KW Signal; MHC.
 FT SIGNAL 1 24
 FT NON_TER 138 138
 SQ SEQUENCE 138 AA; 15610 MW; B8417FA0 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 138;
 Best Local Similarity 90.0%; Pred. No. 0.19;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RNLRLRLRY 10
 ||||| |||
 Db 99 RNLRLRLRY 108

RESULT 6
 O19770 PRELIMINARY; PRT: 172 AA.
 AC O19770;
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)
 DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
 DT 01-NOV-1998 (T-EMBLrel. 08, Last annotation update)
 DE MHC CLASS I HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.

RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,
 RA BEUCHANDRA S., JUJI T., TOKUNAGA K.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U90421; AAB50144.1; -;
 DR EMBL; U90420; AAB50144.1; JOINED.
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 172 172
 SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 172;
 Best Local Similarity 90.0%; Pred. No. 0.24;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RENLRILLY 10
 ||||| |||
 Db 68 RENLRILLY 77

RESULT 7

O19774 PRELIMINARY; PRT; 172 AA.
 AC O19774;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE MHC CLASS I HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,
 RA BEUCHANDRA S., BLASZYK R., GROSSE-WILDE H.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U90423; AAB50145.1; -;
 DR EMBL; U90422; AAB50145.1; JOINED.
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 172 172
 SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 172;
 Best Local Similarity 90.0%; Pred. No. 0.24;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RENLRILLY 10
 ||||| |||
 Db 68 RENLRILLY 77

RESULT 8

O19775 PRELIMINARY; PRT; 172 AA.
 AC O19775;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE MHC CLASS I HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,
 RA BEUCHANDRA S.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; U90425; AAB50146.1; -;
 DR EMBL; U90424; AAB50146.1; JOINED.
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 172 172
 SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 172;
 Best Local Similarity 90.0%; Pred. No. 0.24;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RENLRILLY 10
 ||||| |||
 Db 68 RENLRILLY 77

RESULT 9

O19780 PRELIMINARY; PRT; 172 AA.
 AC O19780;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE MHC CLASS I HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,
 RA BEUCHANDRA S., JUJI T., TOKUNAGA K.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U90419; AAB50143.1; -;
 DR EMBL; U90418; AAB50143.1; JOINED.
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 172 172
 SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 172;
 Best Local Similarity 90.0%; Pred. No. 0.24;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RENLRILLY 10
 ||||| |||
 Db 68 RENLRILLY 77

RESULT 10

O95364 PRELIMINARY; PRT; 172 AA.
 AC O95364;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE MHC HLA-B*51 PROTEIN (FRAGMENT).
 GN HLA-B*51FA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA BLASZYK R.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X96473; CAA65327.1; -;
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1

```

FT NON_TER 172 172
SQ SEQUENCE 172 AA; 19942 MW; 1A73E47D CRC32;

Query Match 89.8%; Score 44; DB 7; Length 172;
Best Local Similarity 90.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENRILLRY 10
   ||||| |||
Db 65 RENLRALRY 74

RESULT 11
O19771 PRELIMINARY; PRT; 172 AA.
AC O19771
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D.,
RA SIRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90612; AAB50151.1; -.
DR EMBL; U90611; AAB50151.1; JOINED.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 20026 MW; 4D9A1043 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 172;
Best Local Similarity 90.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENRILLRY 10
   ||||| |||
Db 68 RENLRALRY 77

RESULT 12
O19772 PRELIMINARY; PRT; 172 AA.
AC O19772
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D.,
RA SIRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90614; AAB50244.1; -.
DR EMBL; U90613; AAB50244.1; JOINED.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 20026 MW; 4D9A1043 CRC32;

```

```

Query Match 89.8%; Score 44; DB 7; Length 172;
Best Local Similarity 90.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENRILLRY 10
   ||||| |||
Db 68 RENLRALRY 77

RESULT 13
O19773 PRELIMINARY; PRT; 172 AA.
AC O19773
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D.,
RA SIRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90616; AAB50245.1; -.
DR EMBL; U90615; AAB50245.1; JOINED.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 20052 MW; F6214671 CRC32;

```

```

Query Match 89.8%; Score 44; DB 7; Length 172;
Best Local Similarity 90.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RENRILLRY 10
   ||||| |||
Db 68 RENLRALRY 77

RESULT 14
O29694 PRELIMINARY; PRT; 175 AA.
AC O29694
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE MHC CLASS I HLA-B ANTIGEN (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BLOOD;
RA PETERSDORF E.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U28759; AAB60367.1; -.
DR HSSP; P10318; 1ROG.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
FT VARIANT 67 67 Y -> D.
FT VARIANT 73 73 I -> T.
FT NON_TER 175 175
SQ SEQUENCE 175 AA; 20332 MW; 83A0C5C3 CRC32;

```


Query Match 89.8%; Score 44; DB 7; Length 175;
 Best Local Similarity 90.0%; Pred. No. 0.24;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RNLRLRLRY 10
 Db 68 RNLRLRLRY 77

RESULT 15

O19607
 ID O19607 PRELIMINARY; PRT; 180 AA.
 AC O19607;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE MHC CLASS I HLA-A (FRAGMENT).
 GN HLA-A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
 RA RUNGRONG E., BEICHANDRA S.;
 RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF030920; AAB87056.1; -.
 DR EMBL; AF030919; AAB87056.1; JOINED.
 DR HSSP; P01891; ITMC.
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 180 180
 SQ SEQUENCE 180 AA; 20811 MW; CECC3537 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 180;
 Best Local Similarity 90.0%; Pred. No. 0.25;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RNLRLRLRY 10
 Db 74 RNLRLRLRY 83

Search completed: February 8, 2000, 13:17:37
 Job time: 32486 sec

THIS PAGE BLANK (USPTO)

OM of: US-08-653-294-10 to: N_Geneseq_36:* out_format : pfs
Date: Feb 8, 2000 1:27 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn1/1/USPTO.spool/US08653294/runat_04022000_160701_15807/app_query.fasta.1
-DB=N_Geneseq_36 -QFMT=fastap -SURFLX=ring -GAPOP=12.000
-CAPEXT=4.000 -MINMATCH=0.100 -LOPCL=0.000 -LOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosome2
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-ALIGN=15 -MODE=LOCAL -OUTENT=pfs -NORM=ext -MINLEN=0
-MAXLEN=1000000 -USER=US08653294 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT
-THREADS=1

Search information block:

Query: US-08-653-294-10
Query length: 10
Database: N_Geneseq_36:*
Database sequences: 311585
Database length: 125096042
Search time (sec): 590.520000

score_list:

| Sequence | Strid | Orig | zScore | Escore | Len | Documentation |
|---------------------|-------|-------|--------|---------|-------|----------------------------------|
| N_Geneseq_36:Q29167 | + | 44.00 | 157.54 | 0.3563 | 270 | HLA-Bw 52 exon 2 alpha-1 domain |
| N_Geneseq_36:Q01834 | + | 44.00 | 145.42 | 1.69 | 1086 | Sequence encoding HLA-B51 anti |
| N_Geneseq_36:Q01822 | + | 44.00 | 145.42 | 1.69 | 1086 | Sequence encoding HLA-Bw52 anti |
| N_Geneseq_36:Q05693 | + | 44.00 | 145.40 | 1.69 | 1089 | HLA-B51 gene for production of |
| N_Geneseq_36:Q05701 | + | 44.00 | 145.40 | 1.69 | 1089 | HLA-Bw52 gene for production of |
| N_Geneseq_36:Q12114 | + | 44.00 | 145.40 | 1.69 | 1089 | HLA-Bw53 exon. HLA-Bw53 gene, |
| N_Geneseq_36:N70935 | + | 39.00 | 127.42 | 16.96 | 1026 | Sequence encoding the human H |
| N_Geneseq_36:N70225 | + | 39.00 | 115.86 | 74.74 | 3874 | Sequence of genomic DNA encodi |
| N_Geneseq_36:T61639 | + | 39.00 | 111.28 | 134.41 | 6553 | HLA B27 consensus sequence. De |
| N_Geneseq_36:T61946 | + | 37.00 | 144.88 | 1.81 | 59 | U7.6 L37 PCR primer for U7.6 var |
| N_Geneseq_36:T12619 | + | 37.00 | 142.23 | 2.54 | 80 | U7.6 L37 PCR primer for U7.6 var |
| N_Geneseq_36:Q69941 | + | 37.00 | 141.91 | 2.64 | 83 | U7.6 L37 PCR primer for U7.6 var |
| N_Geneseq_36:N92449 | + | 37.00 | 125.71 | 21.13 | 534 | Sequence of carcinoembryonic an |
| N_Geneseq_36:Q71567 | + | 37.00 | 114.08 | 93.88 | 2031 | Immunogenic carcinoembryonic a |
| N_Geneseq_36:T36495 | + | 37.00 | 113.96 | 95.33 | 2059 | Carcinoembryonic antigen cDNA |
| N_Geneseq_36:Q82807 | + | 37.00 | 113.80 | 97.29 | 2097 | Carcinoembryonic antigen gene |
| N_Geneseq_36:T33302 | + | 37.00 | 113.30 | 103.69 | 2220 | Carcinoembryonic antigen gene |
| N_Geneseq_36:Q87869 | + | 37.00 | 112.81 | 110.44 | 2349 | H6/CEA expression cassette fro |
| N_Geneseq_36:Q87868 | + | 37.00 | 112.50 | 114.91 | 2434 | H6/CEA expression cassette fro |
| N_Geneseq_36:N81584 | + | 37.00 | 111.16 | 136.45 | 2839 | ILV7 cDNA encoding carcinoemb |
| N_Geneseq_36:N92780 | + | 37.00 | 111.16 | 136.45 | 2839 | cDNA sequence encoding CEA con |
| N_Geneseq_36:Q34352 | + | 37.00 | 111.16 | 136.45 | 2839 | Carcinoembryonic antigen CEA-0 |
| N_Geneseq_36:T46082 | + | 37.00 | 111.16 | 136.45 | 2839 | Carcinoembryonic antigen CEA-0 |
| N_Geneseq_36:V70153 | + | 37.00 | 111.16 | 136.45 | 2839 | CEA protein encoding cDNA. C |
| N_Geneseq_36:N81611 | + | 37.00 | 110.90 | 141.23 | 2928 | Carcinoembryonic antigen gene |
| N_Geneseq_36:T75431 | + | 37.00 | 110.90 | 141.23 | 2928 | Human carcinoembryonic antigen |
| N_Geneseq_36:V82233 | + | 35.00 | 122.79 | 30.70 | 319 | EST clone DX298. New polynucle |
| N_Geneseq_36:Q04288 | + | 35.00 | 113.41 | 102.36 | 938 | Sequence encoding bovine tropho |
| N_Geneseq_36:Q04286 | + | 35.00 | 113.08 | 106.75 | 974 | Sequence encoding bovine tropho |
| N_Geneseq_36:Q04287 | + | 35.00 | 112.42 | 116.22 | 1051 | Sequence encoding bovine tropho |
| N_Geneseq_36:Q04285 | + | 35.00 | 112.37 | 116.96 | 1057 | cDNA clone of sequence encodin |
| N_Geneseq_36:X02073 | + | 35.00 | 99.45 | 613.28 | 4663 | Borrelia burgdorferi polynucle |
| N_Geneseq_36:T95000 | + | 34.00 | 116.64 | 67.62 | 423 | Type II topoisomerase database |
| N_Geneseq_36:Q37687 | + | 34.00 | 101.03 | 500.23 | 2540 | Rat choline kinase gene. Rat C |
| N_Geneseq_36:Q74643 | + | 34.00 | 94.91 | 1.1e+03 | 5136 | Staphylococcus aureus contig S |
| N_Geneseq_36:T11549 | + | 34.00 | 86.44 | 3.3e+03 | 13585 | Tumour rejection antigen pred |
| N_Geneseq_36:X12982 | + | 34.00 | 85.23 | 3.8e+03 | 15614 | Enterococcus faecalis genome |
| N_Geneseq_36:X13060 | + | 34.00 | 78.77 | 8.6e+03 | 32768 | Enterococcus faecalis genome |
| N_Geneseq_36:Q47809 | + | 33.00 | 98.50 | 692.77 | 2223 | Aspartokinase II gene. DNA enc |
| N_Geneseq_36:V11704 | + | 33.00 | 98.48 | 693.81 | 2226 | Enterocin-900 operon. Bacteria |
| N_Geneseq_36:X17728 | + | 33.00 | 98.48 | 693.81 | 2226 | Bacteriocin enterocin 900 form |

N_Geneseq_36:Q25307 + 33.00 94.61 1.1e+03 3473 ! JAK2 encoding DNA. Novel pr
N_Geneseq_36:Q85412 + 33.00 94.23 1.2e+03 3629 ! Murine JAK2 kinase coding s
N_Geneseq_36:V74652 + 33.00 92.25 1.5e+03 4557 ! Staphylococcus aureus conti
N_Geneseq_36:V21209_00 + 33.00 64.53 5.2e+04 110000 ! Methanococcus jannaschi
seq_name: N_Geneseq_36:Q29167
seq_documentation_block:
ID Q29167 standard; DNA; 270 BP.
AC Q29167;
DC 09-MAR-1993 (first entry)
DE HLA-Bw 52 exon 2 alpha-1 domain.
KW Human leukocyte antigen; transgenic; germ cells; somatic cells;
expression; ss.
PN J04091731-A.
PD 25-MAR-1992.
PF 03-AUG-1990; 207329.
PR 03-AUG-1990; JP-207329.
FA (OLYU) OLYMPUS OPTICAL CO.
DR WPI; 92-342893/42.
PT Transgenic non-human mammalian HLA-Bw 52 gene - useful for
analysis of expression of gene structure, and prodn. of
mouse model of human disease
PS Disclosure; Fig 1; 8pp; Japanese.
CC The sequence shows the exon 2 alpha-1-domain of the human leukocyte
antigen-Bw 52 gene. The complete gene may be introduced into non-
human mammals, pref. rat or mouse, or their ancestors at the primary
developmental biological step via transplantation into the zygote or
embryo to generate transgenic non-human mammals incorporating the
HLA-Bw 52 gene in both their germ cells and somatic cells. Transgenic
non-human mammals contg. HLA-Bw 52 are useful for the analysis of
expression of the gene, its structure, and prodn. of mouse models of
human disease. See also Q29166-72.
SQ Sequence 270 BP; 59 A; 88 C; 86 G; 37 T;

alignment_scores:
Quality: 44.00 Length: 10
Ratio: 4.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:

US-08-653-294-10 x Q29167 ..

Align seg 1/1 to: Q29167 from: 1 to: 270

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10

|||||
222 CGAGAGAACTCGGATCGGTCGCTAC 251

seq_name: N_Geneseq_36:Q01834

seq_documentation_block:

ID Q01834 standard; DNA; 1086 BP.

AC Q01834; 1991 (first entry)

DE Sequence encoding HLA-B51 antigen.

KW Probe: HLA class I DNA; immunogen; ss.

OS Homo sapiens.

PN EP354580-A.

PD 14-FEB-1990.

PF 10-AUG-1989.

PR 11-AUG-1988; JP-200758.

FA (OLYU) Olympus Optical Co., Ltd.

PI Kano K, Takiguchi;

DR WPI; 90-046289/07.

PT New DNA for class I human leukocyte antigens and derived probes and

transformed cells, useful for DNA typing, as immunogens etc.

PS Claim 1; Page 11; 23pp; English.

CC The HLA class I DNA can be used as a source of probes for use in DNA

typing. Transformed cells, which are useful as immunogens, can be

obtained by introducing these DNAs into eucaryotic cells.

SQ Sequence 1086 BP; 224 A; 334 C; 356 G; 172 T;

alignment_scores:
 Quality: 44.00 Length: 10
 Ratio: 4.889 Gaps: 0
 Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:

US-08-653-294-10 x Q01834 ..

Align seg 1/1 to: Q01834 from: 1 to: 1086

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
 |||||
 294 CGAGAGAACCTCGGATCGGCTCGCTAC 323

seq_name: N_Geneseq_36:Q01822

seq_documentation_block:

ID Q01822 standard; DNA; 1086 BP.

AC Q01822;

DT 19-MAY-1991 (first entry)

DE Sequence encoding HLA-B*52 antigen.

KW Probe: HLA class I DNA; immunogen; ss.

OS Homo sapiens.

FH Key

FT Key Location/Qualifiers

FT cds 1..1086

FT /*tag= a

PN EP-354580-A.

PD 14-FEB-1990.

PF 10-AUG-1989.

PR 11-AUG-1988; JP-200758.

PA (OLYU) Olympus Optical Co., Ltd.

PI Kano K, Takiguchi;

DR WPI: 90-046289/07.

DR P-PSDB: R03142.

PT New DNA for class I human leucocyte antigens and derived probes and

PT transformed cells, useful for DNA typing, as immunogens etc.

PS Claim 2: p11-12; 23pp: English.

CC The HLA class I DNA can be used as a source of probes for use in DNA

CC typing. Transformed cells, which are useful as immunogens, can be

CC obtained by introducing these DNAs into eucaryotic cells.

SQ Sequence 1086 BP; 223 A; 335 C; 358 G; 170 T;

alignment_scores:
 Quality: 44.00 Length: 10
 Ratio: 4.889 Gaps: 0
 Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:

US-08-653-294-10 x Q01822 ..

Align seg 1/1 to: Q01822 from: 1 to: 1086

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
 |||||
 294 CGAGAGAACCTCGGATCGGCTCGCTAC 323

seq_name: N_Geneseq_36:Q05693

seq_documentation_block:

ID Q05693 standard; DNA; 1089 BP.

AC Q05693;

DT 03-JAN-1991 (first entry)

DE HLA-B*51 gene for production of monoclonal antibodies.

KW Allotype specific monoclonal anti-HLA antibodies; hybridomas;

KW transgenic animals; HLA-B*51 gene; ss.

FH Key

FT Key Location/Qualifiers

FT exon 1..73

FT /*tag= a

FT /number=1

FT 74..343

FT /*tag= b

FT /number=2

FT /*tag= b
 FT /number=2
 FT note="alpha 1-domain"
 FT 344..619
 FT /*tag= c
 FT /number=3
 FT note="alpha 2-domain"
 FT 620..895
 FT /*tag= d
 FT /number=4
 FT note="alpha 3-domain"
 FT 896..1012
 FT /*tag= e
 FT /number=5
 FT 1013..1042
 FT /*tag= f
 FT /number=6
 FT 1043..1089
 FT /*tag= g
 FT /number=7

EP-383183-A.

PD 22-AUG-1990.

PF 07-FEB-1990; 102424.

PR 08-FEB-1989; JP-029313.

PA (OLYU) OLYMPUS OPTICAL KK.

PI Takiguchi M;

DR WPI: 90-255479/34.

PT Allotype specific monoclonal anti- HLA antibodies prodn. - using

PT hybridomas derived from transgenic animals carrying HLA gene and

PT immunised with HLA antigen of different allotype

PS Disclosure: Fig 1 A-G; 20pp: English.

CC The human HLA-B*51 gene was injected into fertilised mouse eggs and

CC then these introduced into the uterus of a pseudo pregnant mouse.

CC The young were tested to ensure incorporation of the gene into the

CC chromosome, and one of them mated 3 times with a normal male to

CC produce 16 young, seven of which carried the HLA-B*51 gene.

CC The transgenic offspring were immunised with HLA antigen.

CC The spleen lymphocytes were fused with myeloma cells. Hybridomas

CC producing antibodies were selected.

CC See also Q05701.

SQ Sequence 1089 BP; 224 A; 335 C; 357 G; 173 T;

alignment_scores:

Quality: 44.00 Length: 10

Ratio: 4.889 Gaps: 0

Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:

US-08-653-294-10 x Q05693 ..

Align seg 1/1 to: Q05693 from: 1 to: 1089

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
 |||||
 295 CGAGAGAACCTCGGATCGGCTCGCTAC 324

seq_name: N_Geneseq_36:Q05701

seq_documentation_block:

ID Q05701 standard; DNA; 1089 BP.

AC Q05701;

DT 03-JAN-1991 (first entry)

DE HLA-B*52 gene for production of monoclonal antibodies.

KW Allotype specific monoclonal anti-HLA antibodies; hybridomas;

KW transgenic animals; HLA-B*52 gene; ss.

FH Key

FT Key Location/Qualifiers

FT exon 1..73

FT /*tag= a

FT /number=1

FT 74..343

FT /*tag= b

FT /number=2

```

FT exon /note="alpha 1-domain"
FT 344..519
FT /*tag= c
FT /number=3
FT /note="alpha 2-domain"
FT 620..895
FT /*tag= d
FT /number=4
FT /note="alpha 3-domain"
FT 896..1012
FT /*tag= e
FT /number=5
FT 1013..1042
FT /*tag= f
FT /number=6
FT 1043..1089
FT /*tag= g
FT /number=7
PN EP-383183-A.
PD 22-AUG-1990.
PF 07-FEB-1990; 102424.
PR 08-FEB-1989; JP-029313.
PA (OLYU ) OLYMPUS OPTICAL KK.
PI Takiguchi M.
DR WPI: 90-255479/34.
PT Allotype specific monoclonal anti- HLA antibodies prodn. - using
PT hybridomas derived from transgenic animals carrying HLA gene and
PT immunised with HLA antigen of different allotype
PS Disclosure; Fig 1 A-G; 20pp; English.
CC The human HLA-Bw52 gene was introduced into mouse L cells and
CC then these cells used to immunise one of the transgenic mice
CC (See 005693)
CC The spleen lymphocytes were fused with myeloma cells (P3x63-Ag8.653).
CC Hybridomas producing antibodies were selected.
SQ Sequence 1089 BP; 223 A; 336 C; 359 G; 171 T;

alignment_scores:
Quality: 44.00 Length: 10
Ratio: 4.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:
US-08-653-294-10 x Q05701 ..
Align seg 1/1 to: Q05701 from: 1 to: 1089

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
|||||:|||||:|||||:|||||:
295 CGAGAGACCTCGGATCGCTCCGCTAC 324

seq_name: N_Geneseq_36:Q12114

seq_documentation_block:
ID Q12114 standard; DNA; 1089 BP.
AC Q12114;1991 (first entry)
DE HLA-Bw53 exon.
KW Human leukocyte antigen; probe; major histocompatibility complex;
KW MHC; class I; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..1089
FT /*tag= a
J03112487-A.
PN 14-MAY-1991.
PD 22-SEP-1989; 247697.
PR 22-SEP-1989; JP-247697.
PA (OLYU ) OLYMPUS OPTICAL KK.
DR WPI: 91-182991/25.
P-PSDB: R12463.
PT HLA-Bw53 gene, DNA probe and transformant cells - used for
PT immunisation, identifying specificity of antiserum etc.

```

```

PS Claim 1: Page 1; 11pp; Japanese.
CC Probes comprising part of the sequence can be used to identify
CC Class I genes. The DNA can be expressed for immunisation of
CC animals and prodn. of monoclonal antibodies specific for the
CC HLA-Bw53 antigen. See also J03112485 and J03112486.
SQ Sequence 1089 BP; 222 A; 337 C; 356 G; 174 T;

alignment_scores:
Quality: 44.00 Length: 10
Ratio: 4.889 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 90.000

alignment_block:
US-08-653-294-10 x Q12114 ..
Align seg 1/1 to: Q12114 from: 1 to: 1089

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
|||||:|||||:|||||:|||||:
295 CGAGAGACCTCGGATCGCTCCGCTAC 324

seq_name: N_Geneseq_36:N70935

seq_documentation_block:
ID N70935 standard; DNA; 1026 BP.
AC N70935;1991 (first entry)
DE Sequence encoding the human histocompatibility antigen HLA B27.
DE Rheumatic disorder; genetic screening; diagnosis;
KW ankylosing spondylitis; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..1026
PN DE3542024-A.
PD 04-JUN-1987.
PF 28-NOV-1985; 542024.
PR 28-NOV-1985; DE-542024.
PR 21-DEC-1985; DE-545576.
PA (BEHW ) BEHRINGER AG.
PI Riethmuller G, Meo T, Weiss E, Szots H;
DR WPI: 87-157893/23.
P-PSDB: P70590.
DR DNA coding for antigen HLA B27 - and diagnostic reagents contg.
PT such DNA, antigen or antibody
PS Claim 2; Page 4; 5pp; German.
CC The DNA may be used as a hybridisation probe for detecting the HLA
CC B27 gene, e.g. for assessing susceptibility to rheumatic disorders
CC such as ankylosis spondylitis, or may be used to transform cells
CC for prodn. of HLA B27. The HLA B27 may be used to detect HLA B27
CC antibody in human serum, or to produce mono- or polyclonal HLA B27
CC antibodies for use in immunoassay.
SQ Sequence 1026 BP; 213 A; 307 C; 344 G; 162 T;

alignment_scores:
Quality: 39.00 Length: 10
Ratio: 4.333 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 80.000

alignment_block:
US-08-653-294-10 x N70935 ..
Align seg 1/1 to: N70935 from: 1 to: 1026

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
|||||:|||||:|||||:|||||:
223 CGAGAGACCTCGGACCTCGCTCCGCTAC 252

seq_name: N_Geneseq_36:N70225

seq_documentation_block:
ID N70225 standard; DNA; 3874 BP.

```

```

AC N70225;
DT 03-APR-1991 (first entry)
DE Sequence of genomic DNA encoding human histocompatibility antigen
DE HLA-B 27.
KW Ankylosing spondylitis; rheumatic disorder; diagnosis; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Intron 518..590
FT Intron /*tag= a
FT Intron 720..989
FT Intron /*tag= b
FT Intron 1090..1506
FT Intron /*tag= c
FT Intron 1932..2257
FT Intron /*tag= d
FT Intron 2450..2566
FT Intron /*tag= e
FT Intron 3009..3041
FT Intron /*tag= f
FT Intron 3148..3191
FT Intron /*tag= g
FT EP-226069-A.
FT PD 24-JUN-1987.
FT PD 21-NOV-1986; 116139.
FT PR 01-JAN-1985; DE-542024.
FT PR 21-DEC-1985; DE-545576.
FT PA (BEHW ) BEHRINGER AG.
FT PI Szots H, Weiss E, Dorner C, Lang M, Meo T, Riethmuller G;
FT DR WPI: 87-171469/25.
FT DR P-PSDB: P70155.
FT PT DNA coding for human histocompatibility antigen HLA-B 27 - useful
FT for diagnosis and antigen and antibody prodn.
FT PS Claim 1: p6; 13pp; German.
FT CC The DNA may be used to detect the HLA-B 27 gene (opt. mutated) in
FT CC human genetic material. The HLA-B 27 may be used to detect anti-HLA-
FT CC B 27 antibodies in human serum. The antibodies may be used to
FT CC determine HLA-B 27 levels in human serum, e.g. for diagnosis of
FT CC rheumatic disorders, esp. ankylosing spondylitis.
FT SQ Sequence 3874 BP; 751 A; 1094 C; 1171 G; 858 T;

alignment_scores:
Quality: 39.00 Length: 10
Ratio: 4.333 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 80.000

alignment_block:
US-08-653-294-10 x N70225 ..
Align seg 1/1 to: N70225 from: 1 to: 3874

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
|||||:||||| |||||
941 CGAGAGACCTGCGACCTGCTCGCTAC 970

seq_name: N_Geneseq_36:T61639

seq_documentation_block:
ID T61639 standard; DNA; 6553 BP.
AC T61639;
DT 05-JUN-1997 (first entry)
DE HLA B27 consensus sequence.
KW HLA B27; seronegative spondylarthropathy; ankylosing spondylitis;
KW Reiter's syndrome; arthritis; acute anterior uveitis; diagnosis;
KW ss; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Intron 3968..6653
FT Intron /*tag= a
FT Intron /*note= "HLA-B27 3' flanking region, downstream of
FT Intron 3' untranslated region"
FT Intron 4112..4556
FT Intron /*tag= b

```

```

FT /note= "3' flanking region diagnostic for genetic
FT predisposition to SNSA"
FT 4270..4556
FT /*tag= b
FT /note= "3' flanking region diagnostic for genetic
FT predisposition to SNSA"
FT misc_difference 4495
FT /*tag= d
FT /*note= "absence of cytosine at this site is
FT indicative of a predisposition to SNSA"
FT WO9709450-A1.
FT PD 13-MAR-1997.
FT PD 16-AUG-1996; U13256.
FT PR 01-SEP-1995; US-522942.
FT PA (CEDA-) CEDARS SINAI MEDICAL CENT.
FT PI Ivan DB;
FT DR WPI: 97-192924/17.
FT PT Detecting pre-disposition to seronegative spondylarthropathies -
FT from the absence of a C residue at a specific position in the
FT 3'-flanking region of the HLA B27 allele
FT PS Claim 1: Page 52-56; 68pp; English.
FT CC Genetic predisposition to seronegative spondylarthropathies (SNSA)
FT is detected by determining the absence of a cytosine nucleotide in
FT the 3' flanking region (see also T61647-48) of an HLA-B gene at a
FT position corresponding to nucleotide 4495 of the HLA-B27 consensus
FT sequence given in T61639. Probes and primers (see also T61640-46)
FT based on this region can be used in diagnostic assays to detect the
FT genetic predisposition to SNSA, and permit the distinction of B27+
FT individuals who are resistant to SNSA from B27+ normal individuals
FT who are susceptible (but as yet unaffected) to such diseases.
FT SQ Sequence 6553 BP; 1443 A; 1619 C; 2017 G; 1474 T;

alignment_scores:
Quality: 39.00 Length: 10
Ratio: 4.333 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 80.000

alignment_block:
US-08-653-294-10 x T61639 ..
Align seg 1/1 to: T61639 from: 1 to: 6553

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
|||||:||||| |||||
1102 CGAGAGACCTGCGACCTGCTCGCTAC 1131

seq_name: N_Geneseq_36:Q69946

seq_documentation_block:
ID Q69946 standard; DNA; 59 BP.
AC Q69946;
DT 26-MAR-1995 (first entry)
DE U7.6 L3' PCR primer for U7.6 variable regions.
KW Marker; antibody; single chain Fv fusion protein; sfv; ss;
KW amplification.
OS synthetic.
PN WO9415642-A.
PD 21-JUL-1994.
PF 07-JAN-1994; U00261.
PR 08-JAN-1993; US-002324.
PA (CREA-) CREATIVE BIOMOLECULES INC.
PA (USSH ) US DEPT HEALTH & HUMAN.
PA SERVICES.
PI George AJT, Huston JS, Segal DM;
DR WPI: 94-248905/30.
DR Delivering agents to target cells - where monospecific binding
PT proteins are administered to a host and bind to target cells,
PT followed by admin. of multivalent antibodies to direct the agents
PT to the target cells
PT Example; Page 29; 80pp; English.
CC The sequence is that of a PCR primer used to isolate the U7.6
CC variable region gene by PCR from murine mRNA.

```

CC See also Q69933-49.
SQ Sequence 59 BP; 7 A; 29 C; 11 G; 12 T;

alignment_scores:
Quality: 37.00 Length: 10
Ratio: 3.700 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 60.000

alignment_block:

US-08-653-294-10 x Q69946 ..

Align seg 1/1 to: Q69946 from: 1 to: 59

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
|||||:|||||:|||||:|||||:|||||:

2 CGCCAGAACCTCCGCTCCTGATCGGCAC 31

seq_name: N_Geneseq_36:T12619

seq_documentation_block:

ID T12619 standard; DNA; 80 BP.

AC T12619;

DT 20-JUL-1996 (first entry)

DE 2C11 scFv VL PCR primer 6.

KW Single-chain Fv; scFv; glycosylation; protein secretion;

KW endoplasmic reticulum; antibody engineering; primer; PCR;

KW polymerase chain reaction; ss.

OS Synthetic.

PN W09605228-A1.

PD 22-FEB-1996.

PF 14-AUG-1995; U10348.

PR 17-AUG-1994; US-292124.

PA (CREA-) CREATIVE BIOMOLECULES INC.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Huston JS, Jost CR, Segal DM;

DR WPI; 96-139645/14.

PT Single-chain Fv molecules with additional glycosylation sites - have

PT increased rates of secretion, decreased antigenicity, modified

PT ligand binding affinity and are protected from proteolytic

PT degradation

PS Example 1: Page 27; 65pp; English.

CC PCR primer 5 (T12618) was designed to introduce a BamHI site

CC into cDNA coding for the VL region of anti-mouse CD3-epsilon

CC chain scFv 2C11. It was used with antisense primer 6 (T12619),

CC which introduces a (GAS)3 linker and NotI site, for the PCR

CC amplification of 2C11 VL DNA. VH cDNA was amplified using

CC primers 7 and 8 (T12620-21). The PCR products were subcloned into

CC pcDNA/AMP and expressed in COS-7 transfectants. Biologically

CC active scFv 2C11 was secreted into the cell culture medium.

SQ Sequence 80 BP; 12 A; 36 C; 14 G; 18 T;

alignment_scores:
Quality: 37.00 Length: 10
Ratio: 3.700 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 60.000

alignment_block:

US-08-653-294-10 x T12619 ..

Align seg 1/1 to: T12619 from: 1 to: 80

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10
|||||:|||||:|||||:|||||:|||||:

26 CGCCAGAACCTCCGCTCCTGATCGGCAC 55

seq_name: N_Geneseq_36:Q69941

seq_documentation_block:

ID Q69941 standard; DNA; 83 BP.

AC Q69941;

DT 26-MAR-1995 (first entry)

DE VK3'AL2 PCR primer for U7.6 variable regions.
KW Marker; antibody; single chain Fv fusion protein; scFv; ss;
KW amplification.
OS Synthetic.
PN W09415642-A.
PD 21-JUL-1994.
PF 07-JAN-1994; U00261.
PR 08-JAN-1993; US-002324.
PA (CREA-) CREATIVE BIOMOLECULES INC.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI George AJT, Huston JS, Segal DM;
DR WPI; 94-248905/30.
PT Delivering agents to target cells - where monospecific binding
PT proteins are administered to a host and bind to target cells,
PT followed by admin. of multivalent antibodies to direct the agents
PT to the target cells
PS Example: Page 29; 80pp; English.
CC The sequence is that of a PCR primer used to isolate the U7.6
CC variable region gene by PCR from murine mRNA.
CC See also Q69933-49.
SQ Sequence 83 BP; 12 A; 39 C; 15 G; 17 T;

alignment_scores:
Quality: 37.00 Length: 10
Ratio: 3.700 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 60.000

alignment_block:

US-08-653-294-10 x Q69941 ..

Align seg 1/1 to: Q69941 from: 1 to: 83

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10

|||||:|||||:|||||:|||||:|||||:

26 CGCCAGAACCTCCGCTCCTGATCGGCAC 55

seq_name: N_Geneseq_36:N92449

seq_documentation_block:

ID N92449 standard; DNA; 534 BP.

AC N92449;

DT 08-MAY-1990 (first entry)

DE Sequence of carcinoembryonic antigen domain III.

KW Carcinoembryonic antigen; Domain III; domain A; domain B.

FH Key Location/Qualifiers

FT cds 1..534

FT /tag= a

FT misc_feature 1..267

FT /tag= b

FT /note="domain A"

FT misc_feature 268..534

FT /tag= c

FT /note="domain B"

FT EP-343946-A.

PN 29-NOV-1989.

PF 24-MAY-1989; 305232.

PR 25-MAY-1988; US-198289.

PA (CITY) City of Hope.

PI Shively JE.

DR WPI; 89-349991/48.

DR P-PSDB; P93499.

PT Carcinoembryonic antigen fragments - used in assays to determine the
PT presence and amt. of the antigen in samples also contg. related antigens.

PS Disclosure: page 4; 15pp; English.

CC CEA fragments can be used in assays to determine the presence and amt. of

CC CEA in samples which also may contain related antigens including its

CC normal cross-reacting antigen or the 128 kD antigen.

SQ Sequence 534 BP; 137 A; 172 C; 110 G; 115 T;

alignment_scores:

Quality: 37.00 Length: 10
Ratio: 4.111 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 70.000

alignment_block:

US-08-653-294-10 x N92449/rev ..

Align seg 1/1 to reverse of: N92449 from: 1 to: 534

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10

429 AAAGAGAACTTGTGTGTGTGTGTGTGTAT 400

seq_name: N_Geneseq_36:Q71567

seq_documentation_block:

ID Q71567 standard; DNA; 2031 BP.
AC Q71567.
DT 11-MAY-1995 (first entry)
DE Carcinoembryonic antigen DNA.
KW Carcinoembryonic antigen; CEA; neoplastic diseases; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT mat_peptide 103..2028
FT /*tag= a
FT EP-618292-A.
PD 05-OCT-1994.
PF 15-MAR-1994; 103986.
PR 25-MAR-1993; EP-810214.
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
PI Mach J, Pelegriin A, Tersikh A;
DR WPI: 94-304461/38.
DR P-PSDB; R60619.
PT Carcinoembryonic antigen (CEA) derives - useful as reagents in
PT immunoassay for diagnosis of neoplastic diseases
PS Claim 4: Page 18; 30pp; English.
CC Q71567 is the DNA sequence which encodes carcinoembryonic antigen (CEA)
CC R60619. CEA is free from cross-reactive CEA-like antigens, it is
CC antigenically indistinguishable from the solution form of CEA shed from
CC tumour cells, and it is devoid of ethanolamine. R60619 can be used in a
CC reagent composition for detecting neoplastic diseases in biological
CC samples, or in an immunoassay process where it can specifically detect
CC the presence of tumour cells in a biological sample e.g. blood.
SQ Sequence 2031 BP; 551 A; 642 C; 417 G; 421 T;

alignment_scores:

Quality: 37.00 Length: 10
Ratio: 4.111 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 70.000

alignment_block:

US-08-653-294-10 x Q71567/rev ..

Align seg 1/1 to reverse of: Q71567 from: 1 to: 2031

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10

1923 AAAGAGAACTTGTGTGTGTGTGTGTAT 1894

seq_name: N_Geneseq_36:T36495

seq_documentation_block:

ID T36495 standard; cDNA; 2059 BP.
AC T36495.
DT 13-NOV-1996 (first entry)
DE immunogenic carcinoembryonic antigen cDNA.
KW Carcinoembryonic antigen; immunogen; breast cancer; lung cancer;
KW colon cancer; therapy; immunotherapy; vaccine; baculovirus; vector;
KW Spodoptera frugiperda; insect; pA9080 ACNPV-CEA; ss.
OS Chimeric Autographa californica nuclear polyhedrosis virus;
OS Chimeric Homo sapiens.
FH Key Location/Qualifiers

FT misc_feature 1..7
FT /*tag= a
FT /function= 7-adenine motif in the polyhedrin mRNA
FT leader
FT cds 18..2009
FT /*tag= b
FT signal_peptide 18..71
FT /*tag= c
FT /function= ACNPV 61k protein signal sequence
FT mat_peptide 72..2006
FT /*tag= d
FT misc_feature 72..80
FT /*tag= e
FT /function= 3 N-terminal residues of the baculovirus
FT (MGS12) not present in human CEA
FT 81..2006
FT /*tag= f
FT /function= mature CEA
FT complement (81..103)
FT /*tag= g
FT /note= "5' primer for CEA amplification"
FT 1985..2006
FT /*tag= g
FT /note= "3' primer for CEA amplification"
FT 2030..2040
FT /*tag= i
FT /function= universal translation termination
FT signal contg. stop codons in all 3
FT reading frames in vector MGS12
PN W09532286-A2.
PD 30-NOV-1995.
PF 19-MAY-1994; US-246981.
PR 20-MAY-1994; US-246981.
PA (MICR-) MICROGENESYS INC.
PI Hackett C, Smith G, Volvovitz F;
DR WPI: 96-020581/02.
DR P-PSDB; R98519.
DR immunogenic carcinoembryonic antigen produced using insect cell
DR baculovirus expression system - useful in cancer therapy
PS Disclosure; Page 47-49; 61pp; English.
CC A portion (T36494) of pA9080 ACNPV-CEA vector codes for recombinant,
CC soluble, immunogenic carcinoembryonic antigen (rCEA) (R98519). The
CC human CEA gene was cloned from colon adenocarcinoma LS174T (ATCC
CC CL 189) cells by PCR (see also T36493-94) and modified for
CC expression in insect cells by replacement of the signal sequence
CC with a 61 kDa baculovirus signal sequence, and deletion of DNA
CC encoding the hydrophobic C-terminal region. The construct was
CC inserted into baculovirus vector pMGS12 contg. a polyhedrin promoter,
CC giving plasmid pA9080 ACNPV-CEA. rCEA was produced at high levels
CC in Sf900+ insect cells grown in serum-free media and isolated to a
CC purity of over 95%. It can be used as an immunogen in humans to
CC protect against breast, lung or colon cancer.
SQ Sequence 2059 BP; 572 A; 623 C; 425 G; 439 T;

alignment_scores:

Quality: 37.00 Length: 10
Ratio: 4.111 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 70.000

alignment_block:

US-08-653-294-10 x T36495/rev ..

Align seg 1/1 to reverse of: T36495 from: 1 to: 2059

1 ArgGluAsnLeuArgIleLeuLeuArgTyr 10

1901 AAAGAGAACTTGTGTGTGTGTGTAT 1872

OM of: US-08-653-294-12 to: EST: * out_format : pfs

Date: Feb 8, 2000 4:02 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn1_1/USPTO_spool/US08653284/runat_04022000_160700_15770/app_query.fasta.1
-DB=EST -QFWT=fastcap -SURFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -XGAPOP=4.500
-XGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -XGAPOP=6.000
-XGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELEX=7.000 -START=1 -NATRI=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -ALIGN=15 -MODE=LOCAL
-OUTFWT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08653294
-NCPU=6 -ICPU=3 -NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-08-653-294-12

Query length: 10

Database: EST:*

Database sequences: 4538634

Database length: 1887831982

Search time (sec): 8553.360000

score_list:

| Sequence | Strd Orig | Zscore | EScore | Len | Documentation |
|--------------------|-----------|--------|--------|--------|---------------|
| gb_est8:C03945 | + | 44.00 | 170.47 | 0.9929 | 232 |
| gb_est10:AA151891 | + | 44.00 | 169.64 | 1.10 | 255 |
| gb_est11:AA263158 | + | 44.00 | 168.73 | 1.24 | 283 |
| gb_est6:D82221 | + | 44.00 | 166.26 | 1.70 | 375 |
| gb_est26:AI359260 | + | 44.00 | 161.87 | 2.99 | 618 |
| gb_est31:AI696864 | + | 44.00 | 160.20 | 3.71 | 748 |
| gb_est10:AA147151 | + | 40.00 | 146.35 | 21.89 | 581 |
| gb_est38:AD036690 | + | 39.00 | 153.07 | 9.25 | 171 |
| gb_est1:R24384 | + | 39.00 | 146.76 | 20.78 | 351 |
| gb_est23:AI124815 | + | 39.00 | 145.31 | 25.02 | 414 |
| gb_est38:AW090252 | + | 38.00 | 141.32 | 41.76 | 413 |
| gb_est6:D82189 | + | 38.00 | 141.27 | 41.99 | 415 |
| gb_gss10:AQ024294 | + | 38.00 | 141.07 | 43.13 | 425 |
| gb_gss11:AQ0317269 | + | 38.00 | 141.02 | 43.36 | 427 |
| gb_gss4:AQ725101 | + | 38.00 | 138.79 | 57.77 | 551 |
| gb_gss14:AQ517553 | + | 38.00 | 138.58 | 59.31 | 564 |
| gb_gss14:AQ512233 | + | 37.00 | 140.34 | 47.31 | 292 |
| gb_gss11:AQ301014 | + | 37.00 | 136.43 | 78.14 | 456 |
| gb_est2:RI3904 | + | 37.00 | 136.37 | 78.72 | 459 |
| gb_gss13:AQ440876 | + | 37.00 | 135.61 | 86.87 | 501 |
| gb_est6:W07747 | + | 37.00 | 135.54 | 87.65 | 505 |
| gb_gss14:AQ570343 | + | 37.00 | 134.60 | 98.87 | 562 |
| gb_gss4:AQ721689 | + | 37.00 | 133.96 | 107.22 | 604 |
| gb_gss7:AQ896303 | + | 37.00 | 131.40 | 148.99 | 809 |
| gb_est21:AA952680 | + | 36.00 | 137.05 | 72.19 | 269 |
| gb_gss12:AQ356093 | + | 36.00 | 137.05 | 72.19 | 269 |
| gb_gss9:AQ101532 | + | 36.00 | 136.27 | 79.79 | 294 |
| gb_gss13:AQ444169 | + | 36.00 | 133.55 | 113.15 | 401 |
| gb_gss4:AQ704919 | + | 36.00 | 132.53 | 128.84 | 450 |
| gb_est37:AI967800 | + | 36.00 | 132.15 | 135.30 | 470 |
| gb_gss12:AQ101394 | + | 36.00 | 131.95 | 138.87 | 481 |
| gb_gss9:AQ165275 | + | 36.00 | 131.54 | 146.37 | 504 |
| gb_gss6:AQ816496 | + | 36.00 | 131.23 | 152.27 | 522 |
| gb_est35:AI831224 | + | 36.00 | 131.21 | 152.59 | 523 |
| gb_est44:AW204764 | + | 36.00 | 131.18 | 153.25 | 525 |
| gb_gss6:AQ882251 | + | 36.00 | 130.29 | 171.77 | 581 |
| gb_gss7:AQ899472 | + | 36.00 | 130.22 | 173.44 | 586 |
| gb_gss6:AQ876890 | + | 36.00 | 130.17 | 174.44 | 589 |
| gb_gss15:AQ625443 | + | 36.00 | 130.04 | 177.44 | 598 |
| gb_est22:AI055656 | + | 36.00 | 129.46 | 191.20 | 639 |
| gb_est30:AU056838 | + | 36.00 | 128.64 | 212.20 | 701 |
| gb_gss1:CNS00L2E | + | 36.00 | 125.65 | 311.56 | 986 |

gb_gss1:CNS00CEB + 36.00 125.52 316.90 1001 ! AL058765 Drosophila melanog
gb_est12:AA298855 - 35.00 132.56 128.43 284 ! AA298855 EST11529 Uterus Hom
gb_est43:AV337956 + 35.00 132.11 136.09 299 ! AV337956 AV337956 RIKEN full

seq_name: gb_est8:C03945

seq_documentation_block:

LOCUS C03945 232 bp mRNA EST 30-JUL-1996
DEFINITION C03945 Human heart cDNA (YNakamura) Homo sapiens cDNA clone
3NHC2454, mRNA sequence.

ACCESSION C03945 GI:1467196

VERSION EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 232)

AUTHORS Tanaka,T., Ogiwara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and Nakamura,Y.

TITLE Construction of a normalized directionally cloned cDNA library from

adult heart and analysis of 3040 clones by partial sequencing

JOURNAL Genomics 35 (1), 231-235 (1996)

MEDLINE 96299762

COMMENT On Oct 24, 1995 this sequence version replaced gi:1040105.

Contact: Yusuke Nakamura

Institute of Medical Science

University of Tokyo

4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan

Tel: 81-3-5449-5372

Fax: 81-3-5449-5433

Email: yusuke@ims.u-tokyo.ac.jp.

Location/Qualifiers

FEATURES

Source

1..232

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="3NHC2454"

/clone_lib="Human heart cDNA (YNakamura)"

/dev_stage="adult"

/note="Organ: heart; normalized directionally cloned cDNA

from adult heart"

BASE COUNT 55 a 77 c 68 g 32 t

ORIGIN

alignment_scores:

Quality: 44.00 Length: 10

Ratio: 4.400 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 90.000

alignment_block:

US-08-653-294-12 x C03945 ..

Align seg 1/1 to: C03945 from: 1 to: 232

1 ArgGluAspLeuArgIleAlaLeuArgTyr 10

|||||:|||||

40 CGAGAACCTCGGATCGCGTCCGCTAC 69

seq_name: gb_est10:AA151891

seq_documentation_block:

LOCUS AA151891 255 bp mRNA EST 10-DEC-1996

DEFINITION AA151891 Homo sapiens colon (#937204) Homo sapiens cDNA clone

IMAGE:566435 5' similar to gb:M15497.cds1 HLA CLASS I

HISTOCOMPATIBILITY ANTIGEN, A*2401 (HUMAN);, mRNA

sequence.

ACCESSION AA151891

VERSION AA151891.1 GI:1720754

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

REFERENCE
AUTHORS
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 255)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissole, S., Dietrich, N., DuBuque, T., Favell, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.,
and Marra, M.

TITLE
JOURNAL
MEDLINE
COMMENT
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

On May 8, 1995 this sequence version replaced gi:800234.

Contact: Willson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Trace considered overall poor quality

Seq primer: -28W13 rev2 from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1. .255

/organism="Homo sapiens"

/db_xref="GDB:459088"

/db_xref="taxon:9606"

/clone="IMAGE:566435"

/lab_host="SOLR cells (kanamycin resistant)"

/note="Organ: colon; Vector: pBluescript SK-; Site: 1:

ECORI; Site: 2: XhoI; Cloned unidirectionally. Primer:

Oligo dT. T-84 colonic epithelial cell line. Average

insert size: 1.0 kb; Uni-ZAP XR vector; -5' adaptor

sequence: 5' GAATTCGACGAG 3' -3' adaptor sequence: 5'

CTCGAGTTTTTTTTTTTTTTT 3"

BASE COUNT 57 a 70 c 75 g 44 t 9 others

ORIGIN

alignment_scores:
Quality: 44.00 Length: 10
Ratio: 4.400 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 90.000

alignment_block:

US-08-653-294-12 x AA151891 ..

Align seg 1/1 to: AA151891 from: 1 to: 255

1 ArgGluAspLeuArgIleAlaLeuArgTyr 10

|||||:|||||:|||||:|||||:|||||:|||||

77 CGAGAGAACCTCGGATCGCTCCGCTAC 106

seq_name: gb_est11:AA263158

seq_documentation_block:

LOCUS AA263158 283 bp mRNA EST 02-JUL-1998

DEFINITION PM0534 KGI-a Lambda Zap Express cDNA library Homo sapiens cDNA 5',

mRNA sequence.

AA263158

VERSION AA263158.1 GI:18998964

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 283)

CAudio, J.O., Liew, C.C., Dempsey, A.A., Cukerman, E., Stewart, A.K.,

Na, E., Atkins, H.I., Iscove, N.N. and Hawley, R.G.

Identification of sequence-tagged transcripts differentially

JOURNAL
MEDLINE
COMMENT

expressed within the human hematopoietic hierarchy

Genomics 50 (1), 44-52 (1998)

98292493

On May 5, 1995 this sequence version replaced gi:797810.

Contact: Hawley RG

Oncology Research Laboratories

The Toronto Hospital

CRCS-424, 67 College St., Toronto, Ontario M5G 2M1, Canada

Tel: 416 3403834

Fax: 416 3403453

Email: r.hawley@utoronto.ca

Similar to M58636 MHC class I HLA-B* gene. Clone was randomly

picked from KGIa primary library.

Seq primer: 5' GAAATTAACCTCATTAAAGG 3'

High quality sequence stop: 283.

Location/Qualifiers

1. .283

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="KGI-a Lambda Zap Express cDNA library"

/cell_type="promyeloblast"

/cell_line="KGI-a"

/note="Vector: Lambda Zap Express (Stratagene); Site: 1:

ECORI; Site: 2: XhoI; Unidirectional cloning sites:

ECORI-XhoI. mRNA was purified from KGI-a cell line, cDNA

was synthesized using an XhoI-OligodT linker primer. EORI

adaptors were ligated, followed by digestion with XhoI for

directional cloning into predigested Lambda Zap Express"

BASE COUNT 64 a 91 c 88 g 40 t

ORIGIN

alignment_scores:

Quality: 44.00 Length: 10

Ratio: 4.400 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 90.000

alignment_block:

US-08-653-294-12 x AA263158 ..

Align seg 1/1 to: AA263158 from: 1 to: 283

1 ArgGluAspLeuArgIleAlaLeuArgTyr 10

|||||:|||||:|||||:|||||:|||||:|||||

120 CGAGAGAACCTCGGATCGCTCCGCTAC 149

seq_name: gb_est6:D82221

seq_documentation_block:

LOCUS D82221 375 bp mRNA EST 09-FEB-1996

DEFINITION HUMHBC4626 Human pancreatic islet Homo sapiens cDNA similar to

HLA-B, mRNA sequence.

ACCESSION D82221

VERSION D82221.1 GI:1183739

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 375)

AUTHORS Takeda, J.

TITLE Human pancreatic islet ESTs

JOURNAL Unpublished (1995)

COMMENT On Apr 14, 1993 this sequence version replaced gi:785255.

Contact: Jun Takeda

Institute for Molecular and Cellular Regulation, Gunma University

3-39-15 Showa-machi, Maebashi Gunma 371, Japan

Tel: 272-20-8856

Fax: 272-20-8896

Email: jtakeda@sb.gunma-u.ac.jp.

Location/Qualifiers

1. .375

/organism="Homo sapiens"

source

Percent Similarity: 100.000 Percent Identity: 90.000

alignment_block:

US-08-653-294-12 x AT696864 ..

Align seg 1/1 to: AT696864 from: 1 to: 748

1 ArgGluspleuArgileAlaLeuArgTyr 10

|||||:|||||:|||||:|||||:|||||

137 CGAGAGACCTCGGATCGCGCTCCGCTAC 166

seq_name: gb_est10:AA147151

seq_documentation_block:

LOCUS AA147151 581 bp mRNA EST 05-DEC-1996
DEFINITION Z032d06.r1 Stratagene colon (#937204) Homo sapiens cDNA clone
IMAGE:588587 5', similar to gb:M64740 HLA CLASS I HISTOCOMPATIBILITY
ANTIGEN, A-24(A-9) A*2402 ALPHA (HUMAN);, mRNA sequence.

ACCESSION AA147151

VERSION AA147151

KEYWORDS EST

SOURCE human

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 581)

AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags

JOURNAL Genome Res. 6 (9), 807-828 (1996)

MEDLINE 97044478

COMMENT On Sep 12, 1996 this sequence version replaced gi:1393699.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28M3 rev2 from Amersham

High quality sequence stop: 272.

FEATURES

Location/Qualifiers

1..581

/organism="Homo sapiens"

/db_xref="GDB:4620889"

/db_xref="taxon:9606"

/clone="IMAGE:588587"

/clone_lib="Stratagene colon (#937204)"

/lab_host="SOLR cells (kanamycin resistant)"

/note="Organ: colon; Vector: pBluescript SK-; Site_1:

EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:

Oligo dt. T-84 colonic epithelial cell line. Average

insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor

sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5'

CTCGAGTTTTTTTTTTT 3'

BASE COUNT 134 a 162 c 185 g 85 t 15 others

ORIGIN

alignment_scores:

Quality: 40.00 Length: 10

Ratio: 4.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 80.000

alignment_block:

US-08-653-294-12 x AA147151 ..

Align seg 1/1 to: AA147151 from: 1 to: 581

1 ArgGluspleuArgileAlaLeuArgTyr 10

|||||:|||||:|||||:|||||:|||||

152 CGAGAGACCTCGGATCGCGCTCCGCTAC 181

seq_name: gb_est38:AL036690

seq_documentation_block:

LOCUS AL036690 171 bp mRNA EST 27-SEP-1999
DEFINITION DXFZP564D2463_r1 584 (synonym: hfbr2) Homo sapiens cDNA clone
IMAGE:588587 5', mRNA sequence.

ACCESSION AL036690

VERSION AL036690

KEYWORDS EST

SOURCE human

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 171)

AUTHORS Dueterhoeft, A., Lauber, J., Mewes, H.W., Gassenhuber, J. and

Wiemann, S.

TITLE EST (Dueterhoeft, et al.)

JOURNAL Unpublished (1999)

COMMENT On Jul 7, 1999 this sequence version replaced gi:5866258.

Contact: Dueterhoeft A

MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by Qiagen within the cDNA sequencing consortium of the

German Genome Project.

No sl sequence available.

This clone is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

Location/Qualifiers

1..171

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="DXFZP564D2463"

/clone_lib="564 (synonym: hfbr2)"

/tissue_type="brain"

/dev_stage="fetal"

/lab_host="Xl-2blue"

/note="Vector: pAMP1; Site_1: NotI; Site_2: SalI"

BASE COUNT 36 a 53 c 60 g 22 t

ORIGIN

alignment_scores:

Quality: 39.00 Length: 10

Ratio: 4.875 Gaps: 0

Percent Similarity: 80.000 Percent Identity: 80.000

alignment_block:

US-08-653-294-12 x AL036690 ..

Align seg 1/1 to: AL036690 from: 1 to: 171

1 ArgGluspleuArgileAlaLeuArgTyr 10

|||||:|||||:|||||:|||||:|||||

62 CGAGAGACCTCGGACCTCGCTCCGCTAC 91

seq_name: gb_est1:T24384

seq_documentation_block:

LOCUS T24384 351 bp mRNA EST 28-JUL-1995

DEFINITION crs1519 lambdaZAPST Ricinus communis cDNA clone pcrl519, mRNA

sequence.

ACCESSION T24384

VERSION T24384.1 GI:689202

```

KEYWORDS
SOURCE Ricinus communis
ORGANISM

REFERENCE
1 (bases 1 to 351)
vandeLoe,F.J., Turner,S. and Somerville,C.
AUTHORS Expressed sequence tags from developing castor seeds
TITLE Plant Physiol. 108, 1141-1150 (1995)
JOURNAL Contact: Somerville CR
COMMENT Carnegie Institution
Carnegie Institution, 290 Panama St, Stanford, CA 94305
Tel: 4153251521
Email: crs@andrew.stanford.edu
Seq primer: T3.

FEATURES
Source
Location/Qualifiers
1..351
/organism="Ricinus communis"
/strain="Baker 296"
/db_xref="taxon:3988"
/clone="pcrs1519"
/clone_lib="lambdaZAPST"
Note:Vector: lambdaZAPII; Site_1: EcoRI; Site_2: XhoI;
Poly(A)+ RNA was purified from developing stage III to
stage V (Greenwood & Bewley, Can. J. Bot. 60:1751-1760,
1982) endosperm plus embryo of immature castor fruits.
CDNA was synthesized and cloned into lambdaZAPII according
to the instructions of the manufacturer (Stratagene);
synthesis was primed from the poly(A) tail, and cloned
directionally into XhoI (3') and EcoRI (5') sites. In few
cases, sequence data indicated that this directionality
was reversed. Partial cDNA clones predominate.

BASE COUNT 104 a 61 c 77 g 102 t 7 others
ORIGIN

alignment_scores
Quality: 39.00 Length: 10
Ratio: 4.333 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 80.000

alignment_block
US-08-653-294-12 x T24384/rev ..
Align seg 1/1 to: T24384 from: 1 to: 351

1 ArgGluAspLeuArgIleAlaLeuArgTyr 10
|||||:|||||:|||||:|||||:|||||:
303 AGACAAAATTGAGGATTGCCTAAGATAC 274

seq_name: gb_est23:AI124815

seq_documentation_block:
LOCUS AI124815 414 bp mRNA 11-SEP-1998
DEFINITION am56e06.x1 Johnston frontal cortex Homo sapiens cDNA clone
IMAGE:1539586 3' similar to gb:M24038.cdsl HLA CLASS I
HISTOCOMPATIBILITY ANTIGEN, BW-44(B-12) B*4402 (HUMAN); contains
MER22.t3 TAKI repetitive element ;, mRNA sequence.
ACCESSION AI124815
VERSION AI124815.1 GI:35933329
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 414)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le.N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Scheilberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project

JOURNAL
COMMENT
Unpublished (1997)
On Jan 17, 1998 this sequence version replaced gi:1899887.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amersham.

FEATURES
Source
Location/Qualifiers
1..414
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1539586"
/clone_lib="Johnston frontal cortex"
/sex="male"
/tissue_type="pooled frontal lobe"
/dev_stage="adult"
/lab_host="SOLR (kanamycin resistant)"
Note="Organ: brain; Vector: Bluescript SK-; Site_1:
EcoRI; Stanley Neuropathology Consortium
(www.stanleylab.org) brains S-58, S-65, S-67, S-78.
Random + oligo-dT primed into EcoRI site of ZAP II Vector.
Mass excised. Avg insert length 1.9kb. Custom library
provided by Dr. Nancy Johnston [(410) 614-3918,
nlj@welchlink.welch.jhu.edu].

BASE COUNT 80 a 140 c 136 g 58 t
ORIGIN

alignment_scores
Quality: 39.00 Length: 10
Ratio: 4.333 Gaps: 0
Percent Similarity: 90.000 Percent Identity: 80.000

alignment_block
US-08-653-294-12 x AI124815 ..
Align seg 1/1 to: AI124815 from: 1 to: 414

1 ArgGluAspLeuArgIleAlaLeuArgTyr 10
|||||:|||||:|||||:|||||:|||||:
284 CGAGAGAACTGCGACCGCGCTCCGCTAC 313

seq_name: gb_est38:AW090252

seq_documentation_block:
LOCUS AW090252 413 bp mRNA 15-OCT-1999
DEFINITION xc85g06.x1 NCI_CGAP_Brn35 Homo sapiens cDNA clone IMAGE:2591098 3',
mRNA sequence.
ACCESSION AW090252
VERSION AW090252.1 GI:6047596
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 413)
AUTHORS NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/ETGAP), Tumor Gene Index
JOURNAL
COMMENT
Unpublished (1998)
On Jul 8, 1999 this sequence version replaced gi:5422650.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.

```

DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Seq primer: -40UP from Gibco
 High quality sequence stop: 401.
 Location/Qualifiers
 1. 413

FEATURES

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2591098"
 /clone_lib="NCI-CGAP_Brn35"
 /tissue_type="tumor, 5 pooled (see description)"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: Salt;
 Site_2: NCI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.33 kb. Tumor types include:
 meningioma, oligodendroglioma, astrocytoma (grade II),
 medulloblastoma, astrocytoma (grade IV). Life Technologies
 catalog #: 11544-012"

BASE COUNT 100 a 107 c 84 g 122 t
 ORIGIN

alignment_scores:

Quality: 38.00 Length: 9
 Ratio: 4.222 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 77.778

alignment_block:

US-08-653-294-12 x AW090252/rev ..
 Align seg 1/1 to reverse of: AW090252 from: 1 to: 413

2 GluAspLeuArgIleAlaLeuArgTyr 10

|||||:|||||:|||||:|||||:|||||
 277 GAGGATGTTAGGATTTCACTCCGTTAC 251

seq_name: gb_est6:D82189

seq_documentation_block:

LOCUS D82189 415 bp mRNA EST 09-FEB-1996
 DEFINITION HUMHBC4524 Human pancreatic islet Homo sapiens cDNA similar to
 HLA-B, mRNA sequence.

ACCESSION D82189

VERSION D82189.1 GI:1183662

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 415)

AUTHORS Takeda,J.

TITLE Human pancreatic islet ESTs

JOURNAL Unpublished (1995)

COMMENT On Apr 14, 1993 this sequence version replaced gi:785206.

Contact: Jun Takeda

Institute for Molecular and Cellular Regulation, Gunma University

3-39-15 Showa-machi, Maebashi Gunma 371, Japan

Tel: 272-20-8856

Fax: 272-20-8896

Email: jtakeda@sb.gunma-u.ac.jp.

Location/Qualifiers

FEATURES

source

1. 415

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="Human pancreatic islet"

/note="vector: Lambda ZAPII; Site_1: Eco RI; Site_2: Xho
 I; mRNA was prepared from normal adult human islets. cDNA
 was directionally synthesized from the Xho I in the vector
 to the EcoRI site. cDNA was size fractionated to remove
 sequences <1000 bp in size."

BASE COUNT 79 a 134 c 133 g 61 t 8 others
 ORIGIN

alignment_scores:

Quality: 38.00 Length: 10
 Ratio: 4.222 Gaps: 0
 Percent Similarity: 90.000 Percent Identity: 80.000

alignment_block:

US-08-653-294-12 x D82189 ..

Align seg 1/1 to: D82189 from: 1 to: 415

1 ArgGluAspLeuArgIleAlaLeuArgTyr 10

|||||:|||||:|||||:|||||:|||||

279 CCAGAGACCTGCGATCGCGTCGNTAC 308

seq_name: gb_gss10:AQ204294

seq_documentation_block:

LOCUS AQ204294 425 bp DNA GSS 17-SEP-1998
 DEFINITION HS_3113_B2_C02_MR_CIT Approved Human Genomic Sperm Library D Homo
 sapiens genomic clone Plate=3113 Col=4 Row=F, genomic survey
 sequence.

ACCESSION AQ204294

VERSION AQ204294.1 GI:3614864

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 425)

AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

COMMENT Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel.: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 3113 row: F column: 4

Class: BAC ends

High quality sequence stop: 425.

Location/Qualifiers

source

1..425

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="Plate-3113 Col=4 Row=F"

/note="lib="CIT Approved Human Genomic Sperm Library D"

/sex="male"

/note="Organ: sperm; Vector: pBelBAC11; BAC Clones in

E-Coli DH10B"

BASE COUNT 116 a 103 c 58 g 147 t 1 others

ORIGIN

alignment_scores:

Quality: 38.00 Length: 9
 Ratio: 4.222 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 77.778

alignment_block:

US-08-653-294-12 x AQ204294/rev ..

Align seg 1/1 to reverse of: AQ204294 from: 1 to: 425

```

1 ArgGluAspLeuArgIleAlaLeuArg 9
||||:|||||:|||||:|||||
161 AGAGATGATTGAGGTGGCTTTACGA 135

seq_name: gb_gss11:AQ317269

seq_documentation_block:
LOCUS      AQ317269      427 bp      DNA      GSS      04-MAY-1999
DEFINITION RP11-1105A4.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-105A4,
            genomic survey sequence.
ACCESSION  AQ317269
VERSION    AQ317269.1 GI:4048520
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 427)
AUTHORS    Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE      Use of human BAC End Sequences for Sequence-Ready Map Building
            Unpublished (1998)
JOURNAL    Contact: Shaying Zhao, William Niernan, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
COMMENT     Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..427
                     /organism="Homo sapiens"
                     /db_xref="GDB:753993"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-105A4"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPC11 Human Male BAC Library"
BASE COUNT          114 a 126 c 92 g 95 t
ORIGIN

alignment_scores:
    Quality: 38.00      Length: 9
    Ratio: 4.222      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 77.778

alignment_block:
US-08-653-294-12 x AQ317269/rev ..
Align seg 1/1 to reverse of: AQ317269 from: 1 to: 427

1 ArgGluAspLeuArgIleAlaLeuArg 9
|||||:|||||:|||||:|||||
31 AGAGAGACTTGGGTAGCCTCAGA 5

seq_name: gb_gss4:AQ725101

seq_documentation_block:
LOCUS      AQ725101      551 bp      DNA      GSS      14-JUL-1999
DEFINITION HS_5392_A2_A05_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=968 Col=10 Row=A, genomic survey sequence.
ACCESSION  AQ725101
VERSION    AQ725101.1 GI:5484770
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 551)
AUTHORS    Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Kellar,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE      Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
            Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
JOURNAL    Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
COMMENT     Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 968 row: A column: 10
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 551.
FEATURES             Location/Qualifiers
     source           1..551
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=968 Col=10 Row=A"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACE3.6; Genomic sequence of BAC ends"
BASE COUNT          120 a 114 c 93 g 218 t 6 others
ORIGIN

alignment_scores:
    Quality: 38.00      Length: 10
    Ratio: 3.800      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 60.000

alignment_block:
US-08-653-294-12 x AQ725101/rev ..
Align seg 1/1 to reverse of: AQ725101 from: 1 to: 551

1 ArgGluAspLeuArgIleAlaLeuArgTyr 10
|||||:|||||:|||||:|||||
190 AGAGAAGACATACAAATGCCATCAGGTAT 161

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 01:29:38 ; Search time 122.56 seconds
(without alignments)
1.933 Million cell updates/sec

Title: US-08-653-294-12
Perfect score: 49
Sequence: 1 REDLRALRY 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 49 | 100.0 | 10 | R83094 | HLA-B2702 CTL modu |
| 2 | 49 | 100.0 | 10 | R95425 | HLA-B2702.75-84(D) |
| 3 | 49 | 100.0 | 10 | W07513 | T-cell modulating |
| 4 | 49 | 100.0 | 10 | W47269 | Immunomodulatory p |
| 5 | 49 | 100.0 | 10 | W33787 | Peptide B2702.75-8 |
| 6 | 44 | 89.8 | 10 | R41208 | Peptide fragment o |
| 7 | 44 | 89.8 | 10 | R83062 | HLA-B2702 CTL modu |
| 8 | 44 | 89.8 | 10 | R95413 | Alpha-helix of HL |
| 9 | 44 | 89.8 | 10 | W07512 | T-cell modulating |
| 10 | 44 | 89.8 | 10 | W47265 | Immunomodulatory p |
| 11 | 44 | 89.8 | 10 | W47271 | Immunomodulatory p |
| 12 | 44 | 89.8 | 10 | W33784 | Peptide B2702.75-8 |
| 13 | 44 | 89.8 | 15 | R92912 | HLA-B2702 CTL modu |
| 14 | 44 | 89.8 | 15 | W33795 | Peptide B2702.70-8 |
| 15 | 44 | 89.8 | 20 | R92907 | HLA-B2702 CTL modu |
| 16 | 44 | 89.8 | 20 | R92908 | HLA-B2702 CTL modu |
| 17 | 44 | 89.8 | 20 | R95428 | HLA-B2702 84-75-84 |
| 18 | 44 | 89.8 | 20 | W33778 | Immunomodulating d |
| 19 | 44 | 89.8 | 20 | W33791 | Peptide B2702.84-7 |
| 20 | 44 | 89.8 | 25 | R41205 | Peptide fragment o |
| 21 | 44 | 89.8 | 25 | R48286 | Peptide fragment o |
| 22 | 44 | 89.8 | 25 | R83090 | HLA-B2702 CTL modu |
| 23 | 44 | 89.8 | 25 | R83093 | HLAB38 CTL modul |
| 24 | 44 | 89.8 | 25 | R95416 | HLA-B2702.60-84. C |
| 25 | 44 | 89.8 | 25 | R95422 | HLAB38.6084. Comps |
| 26 | 44 | 89.8 | 25 | W33794 | Peptide B2702.60-8 |
| 27 | 44 | 89.8 | 184 | Y06801 | Peptide Seq ID No: |
| 28 | 44 | 89.8 | 362 | R03142 | Sequence of HLA-Bw |
| 29 | 44 | 89.8 | 362 | R03144 | Sequence of HLA-B5 |
| 30 | 44 | 89.8 | 362 | R12463 | HLA-Bw53 exon. HLA |
| 31 | 39 | 79.6 | 10 | R41212 | Peptide fragment o |
| 32 | 39 | 79.6 | 10 | R83075 | HLA-B2702 CTL modu |
| 33 | 39 | 79.6 | 10 | R83095 | HLA-B2702 CTL modu |
| 34 | 39 | 79.6 | 10 | R83096 | HLA-B2702 CTL modu |

35 39 79.6 10 1 R95423 HLA-B2705.75-84. C
36 39 79.6 10 1 R95426 HLA-B2702.75-84(T)
37 39 79.6 10 1 W47267 Immunomodulatory p
38 39 79.6 10 1 W33785 Peptide B2705.75-8
39 39 79.6 10 1 W33788 Peptide B2702.75-8
40 39 79.6 10 1 W33789 Peptide B2702.75-8
41 39 79.6 17 1 R71442 Human HLA-B27-(62-
42 39 79.6 17 1 R71443 Human (Phe74)-HLA-
43 39 79.6 20 1 R92909 HLA-B2702 CTL modu
44 39 79.6 20 1 R92910 HLA-B2702 CTL modu
45 39 79.6 20 1 W33792 Peptide B2702.84-7

ALIGNMENTS

RESULT 1

R83094
ID R83094 standard; peptide; 10 AA.
AC R83094;
DT 16-MAY-1996 (first entry)
DE HLA-B2702 CTL modulating Peptide (B2702.75-84(D)).
KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW Immunosuppressant; graft versus host disorder; transplantation; therapy;
KW class I MHC; HLA-B2702.
OS Synthetic.
FN W09526979-A1.
PD 12-OCT-1995.
PF 05-APR-1995; U04349.
PR 05-APR-1994; US-222851.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM, Parham P;
DR WPI: 95-358582/46.
PT Extension of acceptance period of transplants from MHC unmatched
PT donor hosts - using Class I B75-84 MHC antigen of the recipient
PT host
PS Example 14; Page 34; 80pp; English.
CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of
CC class I major histocompatibility complex (MHC) antigens. This sequence
CC corresponds to residues 75-84 of the alpha-1 domain of the class I MHC
CC HLA-B2702. These sequences can be used to extend the period of
CC acceptance by a recipient of a transplant from an MHC unmatched donor.
CC The peptides are administered to a patient in conjunction with
CC a subtherapeutic amount of an immunosuppressant. This is administered
CC the patient for a limited period of time (compared to the lifetime
CC administration for current treatments). The peptides particularly
CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
CC of the patient.
SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDLRALRY 10
| | | | | | | | | |
Db 1 REDLRALRY 10

RESULT 2

R95425
ID R95425 standard; peptide; 10 AA.
AC R95425;
DT 12-NOV-1996 (first entry)
DE HLA-B2702.75-84(D).
KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;
KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
KW cytolysis; antigen presenting cell.
OS Synthetic.
FN Key
FT misc_difference 3 Location/Qualifiers

FT W09513288-A1. /note- "N3D mutation"

PN PD 18-MAY-1995.

PF 10-NOV-1994; U12985.

PR 10-NOV-1993; US-150493.

PS (STRD) UNIV LELAND STANFORD JUNIOR.

PI Clayberger C. Krensky AM;

DR WPI; 95-194027/25.

PT Compens. comprising lymphoid surface membrane proteins - which may

PT inhibit cytolytic activity and differentiation of CTLs.

PS Example; Page 11; 29pp; English.

CC R95413, and R95415-R95431 represent palindromes and fragments of

CC human-leucocyte-associated antigens. This sequence represents the

CC HLA-B2702.75-84(D). These sequences can be used to isolate the protein

CC p74 from a T-cell lysate. p74 is a T-cell surface membrane protein

CC associated with T-cell activation in mammalian T-cells, and is also

CC immunologically cross reactive with the heat shock protein hsc70. p74 is

CC found in a limited number of cell types, but is particularly expressed on

CC B and T cells. p74 can be isolated by lysis of a suitable cell with an

CC amphoteric detergent, and then passed through an affinity column

CC containing a covalently bound HLA-B2702 palindromic peptide.

CC Compositions comprising the extracellular fragment of p74 combined with

CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits

CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate

CC compounds can be screened for their effect on the cytolytic activity of

CC T-cells, by combining them with the extracellular portion of p74 and

CC determining the amount of binding between the candidate compound and p74.

CC Modulation of CTL activity can be inhibited in a cellular composition

CC containing T-cells and antigen presenting cells (APCs). By adding to the

CC mix the extracellular portion of p74, in an amount sufficient to compete

CC with p74 for the binding of the p74 ligand.

CC Sequence 10 AA;

FT Query Match 100.0%; Score 49; DB 1; Length 10;

PN Best Local Similarity 100.0%; Pred. No. 0.00038;

PD Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PF 1 REDLRALRY 10

PR | | | | | | | |

PS 1 REDLRALRY 10

PT 1 REDLRALRY 10

RESULT 3

W07513

ID W07513 standard; peptide; 10 AA.

AC W07513;

DE T-cell modulating peptide #2.

KW T-cell modulator; autoimmune disease; tissue destruction; alpha1-domain;

KW mammal; major histocompatibility complex; MHC class I; antigen; perforin;

KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;

KW rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;

KW thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;

KW autologous target cell; cytokine release; T cell activation; therapy.

OS Synthetic.

PN Buelow R;

PD WPI; 96-518410/51.

PF Treatment of auto-immune disease by admin. of peptide(s) corresp. to

PF major histocompatibility complex antigens - esp. for delaying Onset

PT of clinical symptoms of insulin dependent diabetes by modulating T

PT cell mediated attack on target cells

PS Claim 7; Page 20; 24pp; English.

CC W07512-W07518 represent T-cell modulating peptides that can be used in

CC the method of the invention. These sequences are based on a portion of

CC the generic peptide corresponding to residues 70-91 of the alpha1-domain

CC of the major histocompatibility complex (MHC) class I antigen (see

CC W07510). The method is for affecting the course of an autoimmune disease

CC W07510).

FT Query Match 100.0%; Score 49; DB 1; Length 10;

PN Best Local Similarity 100.0%; Pred. No. 0.00038;

PD Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PF 1 REDLRALRY 10

PR | | | | | | | |

PS 1 REDLRALRY 10

PT 1 REDLRALRY 10

RESULT 5

W33787

ID W33787 standard; peptide; 10 AA.

AC W33787;

DE Peptide B2702.75-84D77 tested for immunomodulating activity.

KW Immunomodulating dimer; immunosuppressant drug; CTL activation;

KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN W09744351-A1.
 PD 27-NOV-1997; U08689.
 PF 22-MAY-1997; US-653294.
 PR (STRD) UNIV LELAND STANFORD JUNIOR.
 PA Beulow R, Clayberger C, Krensky AM;
 DR WPI: 98-086530/08
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Example 1: Page 19; 41pp; English.
 CC Peptides W3784-98 and W3778-9 were assayed for their immunomodulating
 CC activity. A peptide-type compound or variant is claimed which has
 CC immunomodulating activity, including the N-terminal acylated and/or
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where
 CC the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
 CC S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha1 domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.00038;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDLRALRY 10
 |||||
 Db 1 REDLRALRY 10

RESULT 6
 R41208
 ID R41208 standard; peptide; 10 AA.
 AC R41208;
 DT 15-MAR-1994 (first entry)
 DE Peptide fragment of Class I HLA peptide.
 KW Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;
 KW parasitic disease; cytotoxic T lymphocyte; modulation.
 OS Synthetic.
 PN W09317699-A.
 PD 16-SEP-1993.
 PF 25-FEB-1993; U01758.
 PR 02-MAR-1992; US-844716.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger CA, Krensky AM;
 DR WPI: 93-303134/38.
 PT New peptide(s) based on Class I HLA antigen domains - used for
 PT modulating cytotoxic T-lymphocyte activity towards targets
 PS Claim 11; Page 54; 51pp; English.
 CC The peptide is used to modulate cytotoxic T-lymphocyte (CTL)
 CC activity, either by inhibition or stimulation. It can be used
 CC for inhibiting CTL toxicity in transplantations, for inducing CTL
 CC activity in parasitic diseases and neoplasia and in studies on viral
 CC infection. The peptide can also be used for identifying CTLs which
 CC bind to it and removing subsets of CTLs from a T-cell composition.
 CC This peptide sequence is more commonly found within larger peptide
 CC compounds of not more than 30 amino acids in length.

SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.004;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDLRALRY 10
 |||||
 Db 1 REDLRALRY 10

RESULT 7
 R83062
 ID R83062 standard; peptide; 10 AA.
 AC R83062;
 DT 16-MAY-1996 (first entry)
 DE HLA-B*2702 CTL modulating peptide (B2702.75-84).
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW class I MHC; HLA-B*2702.
 OS Synthetic.
 PN W09526979-A1.
 PD 12-OCT-1995.
 PF 05-APR-1995; U04349.
 PR 05-APR-1994; US-222851.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM, Parham P;
 DR WPI: 95-358582/46.
 PT Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B*5-84 MHC antigen of the recipient
 PT host
 PS Claim 15; Page 9; 80pp; English.
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of
 CC class I major histocompatibility complex (MHC) antigens. This sequence
 CC corresponds to residues 75-84 of the alpha-1 domain of the class I MHC
 CC HLA-B*2702. These sequences can be used to extend the period of
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.
 CC The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.004;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDLRALRY 10
 |||||
 Db 1 REDLRALRY 10

RESULT 8
 R95413
 ID R95413 standard; peptide; 10 AA.
 AC R95413;
 DT 12-NOV-1996 (first entry)
 DE Alpha-helix of HLA-B*2702.
 KW HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor;
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;
 KW cytolysis; antigen presenting cell.
 OS Synthetic.
 PN W09513288-A1.
 PD 18-MAY-1995.
 PF 10-NOV-1994; U12985.
 PR 10-NOV-1993; US-150493.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;

DR WPI: 95-194027/25.
PT Compens. comprising lymphoid surface membrane proteins - which may
PS inhibit cytolytic activity and differentiation of CTLs.
PS Example: Page 11: 29pp; English.
CC This sequence represents the alpha-helix of the
CC human-leucocyte-associated antigen B2702 (HLA-B2702). This sequence,
CC epitopes, and palindromes of it (such as R95428) can be used to isolate
CC the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane
CC protein associated with T-cell activation in mammalian T-cells, and is
CC also immunologically cross reactive with the heat shock protein Hsc70.
CC p74 is found in a limited number of cell types, but is particularly
CC expressed on B and T cells. p74 can be isolated by lysis of a suitable
CC cell with an amphoteric detergent, and then passed through an affinity
CC column containing a covalently bound HLA-B2702 palindromic peptide.
CC Compositions comprising the extracellular fragment of p74 combined with
CC HLA-B2702.60-84 (see R95416), induces calcium influx, and inhibits
CC cytotoxic T lymphocyte (CTL) differentiation or cytotoxicity. Candidate
CC compounds can be screened for their effect on the cytolytic activity of
CC T-cells, by combining them with the extracellular portion of p74 and
CC determining the amount of binding between the candidate compound and p74.
CC Modulation of CTL activity can be inhibited in a cellular composition
CC containing T-cells and antigen presenting cells (APCs), by adding to the
CC mix the extracellular portion of p74, in an amount sufficient to compete
CC with p74 for the binding of the p74 ligand.
SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 1; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.004;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDRLALRY 10
||:|||||||
Db 1 RENRLALRY 10
||:|||||||

RESULT 9
W07512 ID W07512 standard; peptide; 10 AA.
AC W07512;
DT 04-AUG-1997 (first entry)
DE T-cell modulating peptide #1.
KW T-cell modulator; autoimmune disease; tissue destruction; alpha-domain;
KW mammal; major histocompatibility complex; MHC class I; antigen; perforin;
KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;
KW rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;
KW thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;
KW autologous target cell; cytokine release; T cell activation; therapy.
OS Synthetic.
PN W09635443-Al.
PD 14-NOV-1996.
PF 05-APR-1996; U04710.
PR 12-MAY-1995; US-440504.
PA (SANG-) SANGSTAT MEDICAL CORP.
PI Buelow R;
DR WPI: 96-518410/51.
PT Treatment of autoimmune disease by admin. of peptide(s) corresp. to
PT major histocompatibility complex antigens - esp. for delaying onset
PT of clinical symptoms of insulin dependent diabetes by modulating T
PT cell mediated attack on target cells
PS Claim 7; Page 20; 24pp; English.
CC W07512-W07518 represent T-cell modulating peptides that can be used in
CC the method of the invention. These sequences are based on a portion of
CC the generic peptide corresponding to residues 70-91 of the alpha-domain
CC of the major histocompatibility complex (MHC) class I antigen (see
CC W07510). The method is for affecting the course of an autoimmune disease
CC involving T-cell mediated destruction of tissue in mammals. These
CC peptides are used especially to treat insulin-dependent diabetes
CC mellitus, preferably being administered during the pre-clinical stage to
CC delay onset of the disease. Other diseases that can be treated are
CC multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus vulgaris,
CC Sjogren's disease, thyroid disease, Hashimoto's thyroiditis, myasthenia
CC gravis, etc. The peptides modulate T-cell mediated attack on autologous

CC target cells, and may also reduce inflammation, swelling, and release of
CC cytokines, perforins, granzymes etc. associated with T cell activation.
SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 1; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.004;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDRLALRY 10
||:|||||||
Db 1 RENRLALRY 10
||:|||||||

RESULT 10
W47265 ID W47265 standard; peptide; 10 AA.
AC W47265;
DT 22-MAY-1998 (first entry)
DE Immunomodulatory peptide.
KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
KW transplant rejection; treatment; autoimmune disease.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 1..10 /note= "at least one of the amino acids is the
FT D-isomer
FT PN W09744052-Al.
PD 27-NOV-1997.
PF 23-APR-1997; U06705.
PR 22-MAY-1996; US-651650.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Clayberger C, Krensky AM;
DR WPI: 98-018220/02.
PT Novel immunomodulatory peptide-type compound - useful for inhibiting
PT transplant rejection
PS Claim 10; Page 36; 41pp; English.
CC The present sequence is an immunomodulatory peptide, which
CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
CC in a pharmaceutical composition together with a subtherapeutic dose
CC of an immunosuppressant, to extend the period of acceptance of a
CC transplant from a major histocompatibility complex (MHC) unmatched
CC donor, i.e. to inhibit transplant rejection. It can also be used in
CC the treatment of autoimmune diseases.
CC Peptides using the D-form amino acids are more effective
CC immunomodulators than their diastereomers or enantiomers.
SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 1; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.004;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDRLALRY 10
||:|||||||
Db 1 RENRLALRY 10
||:|||||||

RESULT 11
W47271 ID W47271 standard; peptide; 10 AA.
AC W47271;
DT 22-MAY-1998 (first entry)
DE Immunomodulatory peptide.
KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
KW transplant rejection; treatment; autoimmune disease.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 1..10 /note= "at least one of the amino acids is the
FT D-isomer
FT FT

PN W09744052-Al.
 PD 27-NOV-1997.
 PF 23-APR-1997; U06705.
 PR 22-MAY-1996; US-651650.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI; 98-018220/02.
 PT Novel immunomodulatory peptide-type compound - useful for inhibiting
 PT transplant rejection
 PS Claim 10; Page 36; 41pp; English.
 CC The present sequence is an immunomodulatory peptide, which
 CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
 CC in a pharmaceutical composition together with a subtherapeutic dose
 CC of an immunosuppressant, to extend the period of acceptance of a
 CC transplant from a major histocompatibility complex (MHC) unmatched
 CC donor, i.e. to inhibit transplant rejection. It can also be used in
 CC the treatment of autoimmune diseases.
 CC peptides using the D-form amino acids are more effective
 CC immunomodulators than their diastereomers or enantiomers.
 SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.004; 1; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 1;

QY 1 REDLRRLRY 10
 |||||
 DB 1 REDLRRLRY 10

RESULT 12

W33784
 ID W33784 standard; peptide; 10 AA.
 AC W33784;
 DE 19-JUN-1998 (first entry)
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN W09744351-Al.
 PD 27-NOV-1997.
 PF 22-MAY-1997; U08689.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI; 98-086530/08.
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Example 1; Page 19; 41pp; English.
 CC Peptides W33784-98 and W33778-9 were assayed for their immunomodulating
 CC activity. A peptide-type compound or variant is claimed which has
 CC immunomodulating activity, including the N-terminal acylated and/or
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where
 CC the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 - E or V; aa77 -
 CC D, S or N; aa79 - R or G; aa80 - I or N; aa81, aa84 - a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 SQ Sequence 10 AA;

Query Match 89.8%; Score 44; DB 1; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.004;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRRLRY 10
 |||||
 DB 1 REDLRRLRY 10

RESULT 13

R29212
 ID R29212 standard; peptide; 15 AA.
 AC R29212;
 DE 16-MAY-1996 (first entry)
 DE HLA-B2702 CTL modulating peptide (B2702.70-84).
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW Class I MHC; HLA-B2702.
 OS Synthetic.
 PN W09526979-Al.
 PD 12-OCT-1995.
 PF 05-APR-1995; U04349.
 PR 03-APR-1994; US-222851.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM, Parham P;
 DR WPI; 95-358582/46.
 PT Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 PT host
 PS Example 15; Page 36; 80pp; English.
 CC R83061-R83085, R83090-R83096 and R92907-R92913 represent fragments of
 CC Class I major histocompatibility complex (MHC) antigens. This sequence
 CC corresponds to residues 70-84 of the alpha-1 domain of the class I MHC
 CC HLA-B2702. These sequences can be used to extend the period of
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.
 CC The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 SQ Sequence 15 AA;

Query Match 89.8%; Score 44; DB 1; Length 15;
 Best Local Similarity 90.0%; Pred. No. 0.0061; 0; Indels 0; Gaps 0;
 Matches 9; Conservative 1; Mismatches 0;

QY 1 REDLRRLRY 10
 |||||
 DB 6 REDLRRLRY 15

RESULT 14

W33795
 ID W33795 standard; peptide; 15 AA.
 AC W33795;
 DE 19-JUN-1998 (first entry)
 DE Peptide B2702.70-84 tested for immunomodulating activity.
 KW Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 OS Synthetic.
 OS Homo sapiens.
 PN W09744351-Al.
 PD 27-NOV-1997.
 PF 22-MAY-1997; U08689.
 PR 24-MAY-1996; US-653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI; 98-086530/08.

PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 PS Example 1: Page 19; 41pp; English.
 CC Peptides W3784-98 and W3778-9 were assayed for their immunomodulating
 CC activity. A peptide-type compound or variant is claimed which has
 CC immunomodulating activity, including the N-terminal acylated and/or
 CC C-terminal amidated or esterified forms of up to 60 amino acids, where
 CC the peptide-type compound comprises the formula: A-B, where A, B =
 CC (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino
 CC acid. The sequence in the brackets may optionally be absent or truncated
 CC at any peptide type bond within the brackets. The compounds comprise
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from
 CC undesirably attacking cells in a host or in vitro. They can also be
 CC used in combination with antigenic peptides or proteins of interest to
 CC activate CTLs. They can also inhibit the proliferation of T cells in
 CC response to anti-CD3. The peptide can be used for preventing rejection
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,
 CC rheumatoid arthritis and lupus erythematosus. The products can also be
 CC used for detection and diagnosis.
 SQ Sequence 15 AA;

QY 1 REDLRALRY 10
 ||:|||||||
 Db 11 RENLRALRY 20

Search completed: February 8, 2000, 01:29:38
 Job time: 1750 sec

Query Match 89.8%; Score 44; DB 1; Length 15;
 Best Local Similarity 90.0%; Pred. No. 0.0061;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
 ||:|||||||
 Db 6 RENLRALRY 15

RESULT 15
 R92907
 ID R92907 standard; peptide: 20 AA.
 AC R92907;
 DT 16-MAY-1996 (first entry)
 DE HLA-B2702 CTL modulating peptide (B2702.84-75/75-84).
 KW Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;
 KW class I MHC; HLA-B2702.
 OS Synthetic.
 PN W09526979-A1.
 PD 12-OCT-1995.
 PF 05-APR-1995; U04349.
 PR 05-APR-1994; US-222851.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM, Parham P;
 DR WPI: 95-358582/46.
 PT Extension of acceptance period of transplants from MHC unmatched
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient
 PT host
 PS Example 15: Page 36; 80pp; English.
 CC R83061-R83085, R83090-R83096 and R92907-R92914 represent fragments of
 CC class I major histocompatibility complex (MHC) antigens. This sequence
 CC is an inverted dimer of residues 75-84 of the alpha-1 domain of the class
 CC I MHC HLA-B2702. These sequences can be used to extend the period of
 CC acceptance by a recipient of a transplant from an MHC unmatched donor.
 CC The peptides are administered to a patient in conjunction with a
 CC subtherapeutic amount of an immunosuppressant. This is administered to
 CC the patient for a limited period of time (compared to the lifetime
 CC administration for current treatments). The peptides particularly
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)
 CC of the patient.
 SQ Sequence 20 AA;

Query Match 89.8%; Score 44; DB 1; Length 20;
 Best Local Similarity 90.0%; Pred. No. 0.0083;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2000, 11:54:23 ; Search time 117.7 Seconds
(without alignments)
4.008 Million cell updates/sec

Title: US-08-653-294-12
Perfect score: 49
Sequence: 1 REDLRALRY 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : PIR_62:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 44 | 89.8 | 273 | 2 | 138509 |
| 2 | 44 | 89.8 | 274 | 2 | 154463 |
| 3 | 44 | 89.8 | 354 | 2 | I59308 |
| 4 | 44 | 89.8 | 354 | 2 | I80168 |
| 5 | 44 | 89.8 | 354 | 2 | I80167 |
| 6 | 44 | 89.8 | 355 | 2 | I80169 |
| 7 | 44 | 89.8 | 355 | 2 | I80171 |
| 8 | 44 | 89.8 | 359 | 1 | HLH012 |
| 9 | 44 | 89.8 | 362 | 1 | HLH088 |
| 10 | 44 | 89.8 | 362 | 2 | B30345 |
| 11 | 44 | 89.8 | 362 | 2 | JH0541 |
| 12 | 44 | 89.8 | 362 | 2 | JH0539 |
| 13 | 44 | 89.8 | 362 | 2 | JH0540 |
| 14 | 44 | 89.8 | 362 | 2 | A45834 |
| 15 | 44 | 89.8 | 362 | 2 | I84486 |
| 16 | 44 | 89.8 | 362 | 2 | I82045 |
| 17 | 44 | 89.8 | 362 | 2 | I84490 |
| 18 | 44 | 89.8 | 362 | 2 | I37521 |
| 19 | 44 | 89.8 | 362 | 2 | A30345 |
| 20 | 44 | 89.8 | 362 | 2 | I59633 |
| 21 | 44 | 89.8 | 362 | 2 | S24434 |
| 22 | 44 | 89.8 | 362 | 2 | I37120 |
| 23 | 44 | 89.8 | 363 | 2 | S07113 |
| 24 | 44 | 89.8 | 363 | 2 | S03537 |
| 25 | 44 | 89.8 | 364 | 2 | D35997 |
| 26 | 44 | 89.8 | 365 | 2 | S77963 |
| 27 | 44 | 89.8 | 365 | 2 | I54416 |
| 28 | 44 | 89.8 | 365 | 2 | I54493 |
| 29 | 43 | 87.8 | 274 | 1 | HLH032 |
| 30 | 43 | 87.8 | 364 | 2 | A35997 |

ALIGNMENTS

RESULT 1

I38509
MHC class I histocompatibility antigen - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 23-Jul-1999
C:Accession: I38509
R:Cereb, N.; Choi, J.W.; Riu, K.Z.; Yang, S.Y.
Tissue Antigens 44, 271-273, 1994
A:Title: HLA-B*5105, a newly identified B51 IEF variant.
A:Reference number: I38509; MUID:95176331
A:Accession: I38509
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-273 <RES>
A:Cross-references: EMBL:U06697; NID:g469544; PIDN:AAA92997.1; PID:g469545
C:Genetics:
A:Gene: GDB:HLA-B
A:Cross-references: GDB:120048; OMIM:142830
A:Map position: 6p21.3-6p21.3
A:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 89.8%; Score 44; DB 2; Length 273;
Best Local Similarity 90.0%; Pred. No. 0.12;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10

Db 74 RENLRALRY 83

RESULT 2

I54463
MHC HLA-B*38 chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
C:Accession: I54463
R:Mueller, C.A.; Engler-Blum, G.; Gekeler, V.; Steiert, I.; Weiss, E.; Schmidt, H.
Immunogenetics 30, 200-207, 1989
A:Title: Genetic and serological heterogeneity of the supertypic HLA-B locus specific
A:Reference number: I54463; MUID:89379286
A:Accession: I54463
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-274 <RES>
A:Cross-references: GB:M29864; NID:g197674; PIDN:AAA36222.1; PID:g187675
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 89.8%; Score 44; DB 2; Length 274;
Best Local Similarity 90.0%; Pred. No. 0.12;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRIALRY 10
||:|||||||
Db 75 RENLRIALRY 84

RESULT 3

I59308

class I histocompatibility antigen - pygmy chimpanzee (fragment)

C:Species: Pan paniscus (pygmy chimpanzee, bonobo)

C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 23-Jul-1999

C:Accession: I59308

R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin

Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994

A:Title: A uniquely high level of recombination at the HLA-B locus.

A:Reference number: I59308; MUID:94286544

A:Accession: I59308

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-354 <RES>

A:Cross-references: EMBL:U05575; NID:g454767; PIDN:AAA50178.1; PID:g454768

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 89.8%; Score 44; DB 2; Length 354;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRIALRY 10

||:|||||||

Db 91 RENLRIALRY 100

RESULT 4

I80168

class I histocompatibility antigen - chimpanzee (fragment)

C:Species: Pan troglodytes (chimpanzee)

C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999

C:Accession: I80168

R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin

Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994

A:Title: A uniquely high level of recombination at the HLA-B locus.

A:Reference number: I59308; MUID:94286544

A:Accession: I80168

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-354 <RES>

A:Cross-references: EMBL:U05579; NID:g454775; PIDN:AAA50182.1; PID:g454776

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 89.8%; Score 44; DB 2; Length 354;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRIALRY 10

||:|||||||

Db 91 RENLRIALRY 100

RESULT 5

I80167

class I histocompatibility antigen - pygmy chimpanzee (fragment)

C:Species: Pan paniscus (pygmy chimpanzee, bonobo)

C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999

C:Accession: I80167

R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin

Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994

A:Title: A uniquely high level of recombination at the HLA-B locus.

A:Reference number: I59308; MUID:94286544

A:Accession: I80167

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-354 <RES>

A:Cross-references: EMBL:U05578; NID:g454773; PIDN:AAA50181.1; PID:g454774
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 89.8%; Score 44; DB 2; Length 354;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRIALRY 10

||:|||||||

Db 91 RENLRIALRY 100

RESULT 6

I80169

class I histocompatibility antigen - chimpanzee (fragment)

C:Species: Pan troglodytes (chimpanzee)

C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999

C:Accession: I80169

R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Wat

Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994

A:Title: A uniquely high level of recombination at the HLA-B locus.

A:Reference number: I59308; MUID:94286544

A:Accession: I80169

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-355 <RES>

A:Cross-references: EMBL:U05580; NID:g454777; PIDN:AAA50183.1; PID:g454778

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 89.8%; Score 44; DB 2; Length 355;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRIALRY 10

||:|||||||

Db 91 RENLRIALRY 100

RESULT 7

I80171

class I histocompatibility antigen - chimpanzee (fragment)

C:Species: Pan troglodytes (chimpanzee)

C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 23-Jul-1999

C:Accession: I80171

R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Wat

Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994

A:Title: A uniquely high level of recombination at the HLA-B locus.

A:Reference number: I59308; MUID:94286544

A:Accession: I80171

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-355 <RES>

A:Cross-references: EMBL:U05582; NID:g454781; PIDN:AAA50185.1; PID:g454782

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 89.8%; Score 44; DB 2; Length 355;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRIALRY 10

||:|||||||

Db 91 RENLRIALRY 100

RESULT 8

HLHUI2

MHC

class I histocompatibility antigen HLA alpha chain precursor (clone pHLA 12.4) -

C:Species: Homo sapiens (man)

C:Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 22-Jun-1999

C:Accession: A02189

R.Malissen, M.; Malissen, B.; Jordan, B.R.
Proc. Natl. Acad. Sci. U.S.A. 79, 893-897, 1982
A:Title: Exon/intron organization and complete nucleotide sequence of an HLA gene.
A:Reference number: A02189; MUID:82151002
A:Accession: A02189
A:Molecule type: DNA
A:Residues: 1-359 <MAL>
A:Cross-references: GB:J00191; GB:V00526; NID:g187600; PIDN:AAA36218.1; PID:g386873
C:Comment: The seven exons correspond approximately to the domain structure of this chain
C:Genetics:
A:Map position: 6p21.3
A:Introns: 22/1; 112/1; 204/1; 296/1; 335/1; 346/1
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplantati
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-359/Product: class I histocompatibility antigen HLA alpha chain #status predicted <
F:22-304/Domain: extracellular #status predicted <EXT>
F:22-111/Domain: alpha-1 <EX1>
F:112-203/Domain: alpha-2 <EX2>
F:121-282/Domain: immunoglobulin homology <IMM>
F:305-329/Domain: transmembrane #status predicted <TMM>
F:335-359/Domain: intracellular #status predicted <INT>
F:107/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:224-280/Disulfide bonds: #status predicted

Query Match 89.8%; Score 44; DB 1; Length 359;
Best Local Similarity 90.0%; Pred. No. 0.16;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDRLALRY 10

Db 96 RENLRLALRY 105

|||||

RESULT 9

HLHUB8
MHC class I histocompatibility antigen HLA-Bw58 alpha chain precursor - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 05-Sep-1997
C:Accession: A23895
R:Ways, J.P.; Coppin, H.L.; Parham, P.
J. Biol. Chem. 260, 11924-11933, 1985
A:Title: The complete primary structure of HLA-Bw58.
A:Reference number: A23895; MUID:86008247
A:Accession: A23895
A:Molecule type: DNA
A:Residues: 1-362 <WAW>
A:Note: The authors translated the codon GCC for residue 349 as Ser
C:Comment: This protein is a subtype of the HLA-B17 family.
C:Genetics:
A:Gene: GDB:HLA-B
A:Cross-references: GDB:120048; OMIM:142830
A:Map position: 6p21.3-6p21.3
A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplantati
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-362/Product: class I histocompatibility antigen HLA-Bw58 alpha chain #status predic
F:25-307/Domain: extracellular #status predicted <EXT>
F:25-114/Domain: alpha-1 <EX1>
F:115-206/Domain: alpha-2 <EX2>
F:220-285/Domain: immunoglobulin homology <IMM>
F:308-331/Domain: transmembrane #status predicted <TMM>
F:332-362/Domain: intracellular #status predicted <INT>
F:110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 89.8%; Score 44; DB 1; Length 362;

Best Local Similarity 90.0%; Pred. No. 0.16;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDRLALRY 10

Db 99 RENLRLALRY 108

|||||

RESULT 10

B30345
MHC class I histocompatibility antigen HLA-Bw52 precursor - human
C:Species: Homo sapiens (man)
C:Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 16-Feb-1997
C:Accession: B30345
R:Hayashi, H.; Ennis, P.D.; Ariga, H.; Salter, R.D.; Parham, P.; Kano, K.; Takiguchi, J.
J. Immunol. 142, 306-311, 1989
A:Title: HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the helical
A:Reference number: A30345; MUID:89080265
A:Accession: B30345
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-362 <HAY>
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: transmembrane protein
F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 89.8%; Score 44; DB 2; Length 362;

Best Local Similarity 90.0%; Pred. No. 0.16;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDRLALRY 10

Db 99 RENLRLALRY 108

|||||

RESULT 11

JH0541
class I histocompatibility antigen Gogo-B0103 heavy chain precursor - lowland gorilla
C:Species: Gorilla gorilla gorilla (lowland gorilla)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C:Accession: JH0541
R:Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.
J. Exp. Med. 174, 1491-1509, 1991
A:Title: Gorilla class I major histocompatibility complex alleles: comparison to huma
A:Reference number: JH0534; MUID:92078860
A:Accession: JH0541
A:Molecule type: DNA
A:Residues: 1-362 <LAW>
A:Cross-references: EMBL:X60254; NID:g22869; PIDN:CAA42806.1; PID:g22870
A:Experimental source: EBV-transformed B cell
C:Genetics:
A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: transmembrane protein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0103 #status
F:25-114/Domain: alpha-1 <AL1>
F:115-206/Domain: alpha-2 <AL2>
F:207-298/Domain: alpha-3 <AL3>
F:220-285/Domain: immunoglobulin homology <IMM>
F:299-362/Domain: intracellular #status predicted <INT>

Query Match 89.8%; Score 44; DB 2; Length 362;

Best Local Similarity 90.0%; Pred. No. 0.16;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDRLALRY 10

Db 99 RENLRLALRY 108

|||||

RESULT 12

JH0539
class I histocompatibility antigen Gogo-B0101 heavy chain precursor - lowland gorilla
C:Species: Gorilla gorilla gorilla (lowland gorilla)

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C:Accession: JH0539
R:Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.
J. Exp. Med. 174, 1491-1509, 1991
A:Title: Gorilla class I major histocompatibility complex alleles: comparison to human
A:Reference number: JH0534; MUID:92078860
A:Accession: JH0539
A:Molecule type: DNA
A:Residues: 1-362 <LAW>
A:Cross-references: EMBL:X60255; NID:g22865; PIDN:CAA42807.1; PID:g22866
A:Experimental source: EBV-transformed B cell
C:Genetics:
A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: transmembrane protein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0101 #status pre
F:25-114/Domain: alpha-1 <AL1>
F:115-206/Domain: alpha-2 <AL2>
F:207-298/Domain: alpha-3 <AL3>
F:220-285/Domain: immunoglobulin homology <IMM>
F:299-362/Domain: intracellular #status predicted <INT>

Query Match 89.8%; Score 44; DB 2; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.16;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
||:|||||||
Db 99 RENDLRALRY 108

RESULT 13
JH0540
class I histocompatibility antigen Gogo-B0102 heavy chain precursor - lowland gorilla
C:Species: Gorilla gorilla gorilla (lowland gorilla)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C:Accession: JH0540
R:Lawlor, D.A.; Warren, E.; Taylor, P.; Parham, P.
J. Exp. Med. 174, 1491-1509, 1991
A:Title: Gorilla class I major histocompatibility complex alleles: comparison to human
A:Reference number: JH0534; MUID:92078860
A:Accession: JH0540
A:Molecule type: DNA
A:Residues: 1-362 <LAW>
A:Cross-references: EMBL:X60693; NID:g22867; PIDN:CAA43101.1; PID:g22868
A:Experimental source: EBV-transformed B cell
C:Genetics:
A:Introns: 25/1; 115/1; 207/1; 299/1; 338/1; 349/1
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: transmembrane protein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-362/Product: class I histocompatibility antigen heavy chain, Gogo-B0102 #status pre
F:25-114/Domain: alpha-1 <AL1>
F:115-206/Domain: alpha-2 <AL2>
F:207-298/Domain: alpha-3 <AL3>
F:220-285/Domain: immunoglobulin homology <IMM>
F:299-362/Domain: intracellular #status predicted <INT>

Query Match 89.8%; Score 44; DB 2; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.16;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
||:|||||||
Db 99 RENDLRALRY 108

RESULT 14
A45834
MHC class I histocompatibility antigen HLA-B53 alpha chain precursor - human

C:Species: Homo sapiens (man)
C:Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 23-Jul-1999
C:Accession: A45834
R:Hayashi, H.; Ooba, T.; Nakayama, S.; Sekimata, M.; Kano, K.; Takiguchi, M.
Immunogenetics 32, 195-199, 1990
A:Title: Allospecificities between HLA-B*53 and HLA-B*35 are generated by substitution
A:Reference number: A45834; MUID:91033941
A:Accession: A45834
A:Molecule type: DNA
A:Residues: 1-362 <HAY>
A:Cross-references: GB:M58636; NID:gl87756; PIDN:AAA36228.1; PID:gl87757; GB:M33574
A:Note: this allele is designated B*5301
C:Genetics:
A:Gene: GDB:HLA-B
A:Cross-references: GDB:120048; OMIM:142830
A:Map position: 6p21.3-6p21.3
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: glycoprotein; heterodimer; transmembrane protein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:220-285/Domain: immunoglobulin homology <IMM>
F:110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 89.8%; Score 44; DB 2; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.16;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
||:|||||||
Db 99 RENDLRALRY 108

RESULT 15
I84486
transmembrane glycoprotein - human
C:Species: Homo sapiens (man)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 23-Jul-1999
C:Accession: I84486
R:Hildebrand, W.H.; Domene, J.D.; Shen, S.Y.; Lau, M.; Terasaki, P.I.; Bunce, M.; Mar
Tissue Antigens 43, 209-218, 1994
A:Title: HLA-B*15: a widespread and diverse family of HLA-B alleles.
A:Reference number: I38421; MUID:94367483
A:Accession: I84486
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-362 <RES>
A:Cross-references: GB:IL15005; NID:g493154; PIDN:AAA56832.1; PID:g493155
C:Genetics:
A:Gene: HLA-B*1513
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: glycoprotein

Query Match 89.8%; Score 44; DB 2; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.16;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
||:|||||||
Db 99 RENDLRALRY 108

Search completed: February 7, 2000, 11:54:24
Job time: 24334 sec

| | | | | |
|----|----------|-----|-----|----------------|
| FT | CARBOHYD | 106 | 106 | BY SIMILARITY. |
| FT | CARBOHYD | 106 | 106 | BY SIMILARITY. |

SQ SEQUENCE 359 AA; 40173 MW; 5395FFC9 CRC32;

Query Match 89.8%; Score 44; DB 1; Length 359;
Best Local Similarity 90.0%; Pred. No. 0.076;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
||:|||||||

Db 95 RENLRALRY 104

RESULT 2

ID 1B01_GORGO STANDARD; PRT; 362 AA.
AC P30379;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE CLASS 1 HISTOCOMPATIBILITY ANTIGEN, GOGO-B0101 ALPHA CHAIN PRECURSOR.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92078860.
RA LAWOR D.A., WARREN E., TAYLOR P., PARHAM P.;
RT "Gorilla class 1 major histocompatibility complex alleles: comparison to human and chimpanzee class I.";
RL J. Exp. Med. 174:1491-1509(1991).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
CC
CC EMBL: X60255; J05339.
CC PIR: JH0539; JH0539.
CC HSSP: P03989; ILSA.
CC PROSITE: PS00290; IG_MHC; 1.
CC PFAM: PF00047; IG_1.
CC MHC I; Transmembrane; Glycoprotein; Signal.
CC SIGNAL 1 24
CC CHAIN 25 362
CC CLASS 1 HISTOCOMPATIBILITY ANTIGEN, GOGO-B0101 ALPHA CHAIN.
CC EXTRACELLULAR ALPHA-1.
CC EXTRACELLULAR ALPHA-2.
CC CONNECTING PEPTIDE.
CC CYTOPLASMIC TAIL.
CC BY SIMILARITY.
CC BY SIMILARITY.
CC BY SIMILARITY.
CC CARBOHYD 110 110
CC SEQUENCE 362 AA; 40170 MW; 2E33E2B8 CRC32;

Query Match 89.8%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.076;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
||:|||||||

Db 99 RENLRALRY 108

RESULT 3

ID 1B02_GORGO STANDARD; PRT; 362 AA.
AC P30380;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE CLASS 1 HISTOCOMPATIBILITY ANTIGEN, GOGO-B0102 ALPHA CHAIN PRECURSOR.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92078860.
RA LAWOR D.A., WARREN E., TAYLOR P., PARHAM P.;
RT "Gorilla class 1 major histocompatibility complex alleles: comparison to human and chimpanzee class I.";
RL J. Exp. Med. 174:1491-1509(1991).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
CC
CC EMBL: X60693; CAA43101.1;
CC PIR: JH0540; JH0540.
CC HSSP: P03989; ILSA.
CC PROSITE: PS00290; IG_MHC; 1.
CC PFAM: PF00047; IG_1.
CC MHC I; Transmembrane; Glycoprotein; Signal.
CC SIGNAL 1 24
CC CHAIN 25 362
CC CLASS 1 HISTOCOMPATIBILITY ANTIGEN, GOGO-B0102 ALPHA CHAIN.
CC EXTRACELLULAR ALPHA-1.
CC EXTRACELLULAR ALPHA-2.
CC CONNECTING PEPTIDE.
CC CYTOPLASMIC TAIL.
CC BY SIMILARITY.
CC BY SIMILARITY.
CC BY SIMILARITY.
CC CARBOHYD 110 110
CC SEQUENCE 362 AA; 40204 MW; 3CF119AD CRC32;

Query Match 89.8%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.076;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
||:|||||||

Db 99 RENLRALRY 108

RESULT 4

ID 1B03_GORGO STANDARD; PRT; 362 AA.
AC P30381;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE CLASS 1 HISTOCOMPATIBILITY ANTIGEN, GOGO-B0103 ALPHA CHAIN PRECURSOR.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92078860.
 RA LAWLOR D.A., WARREN E., TAYLOR P., PARHAM P.;
 RT "Gorilla class I major histocompatibility complex alleles: comparison
 RT to human and chimpanzee class I.";
 RL J. Exp. Med. 174:1491-1509(1991).
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X60254; CAA42806.1; -
 DR PIR; JH0541; JH0541.
 DR HSP; P03989; IHSA.
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; ig; 1.
 DR PFAM; PF00129; MHC_I; 1.
 DR MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT TRANSMEM 309 332
 FT DOMAIN 333 362
 FT DISULFID 125 188
 FT DISULFID 227 283
 FT CARBOHYD 110 110
 SQ SEQUENCE 362 AA; 40248 MW; FEAG6A941 CRC32;
 Query Match 89.8%; Score 44; DB 1; Length 362;
 Best Local Similarity 90.0%; Pred. No. 0.076;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
 DB 99 RENLRALRY 108
 RESULT 5
 ID IB15_HUMAN STANDARD; PRT; 362 AA.
 AC P10317;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-27 B*2702 ALPHA CHAIN
 DE PRECURSOR (B-27K) (B27.2).
 GN HLA-B OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86220133.
 RA SEEMANN G.H.A., REIN R.S., BROWN C.S., PLOEGH H.L.;
 RT "Gene conversion-like mechanisms may generate polymorphism in human
 RT class I genes.";
 RL EMBO J. 5:547-552(1986).
 RN [2]

RP SEQUENCE FROM N.A.
 RA PARHAM P., ARNETT K.L., ADAMS E.J.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 86-107 AND 171-181.
 RX MEDLINE; 86042671.
 RA VEGA M.A., EZQUERRA A., ROJO S., APARICIO P., BRAGADO R.,
 RA LOPEZ DE CASTRO J.A.;
 RT "Structural analysis of an HLA-B27 functional variant: identification
 RT of residues that contribute to the specificity of recognition by
 RT cytolytic T lymphocytes.";
 RT Proc. Natl. Acad. Sci. U.S.A. 82:7394-7398(1985).
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X03664; CAA27301.1; -
 DR EMBL; X03667; CAA27301.1; JOINED.
 DR EMBL; L38504; AAA59724.1; -
 DR PIR; B25092; HLUHUK.
 DR HSP; P03989; IHSA.
 DR MIM; 142830; -
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; ig; 1.
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC I; transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT TRANSMEM 309 332
 FT DOMAIN 333 362
 FT CARBOHYD 110 110
 FT DISULFID 125 188
 FT DISULFID 227 283
 SQ SEQUENCE 362 AA; 40397 MW; 9798F0BB CRC32;
 Query Match 89.8%; Score 44; DB 1; Length 362;
 Best Local Similarity 90.0%; Pred. No. 0.076;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
 DB 99 RENLRALRY 108
 RESULT 6
 ID IB47_HUMAN STANDARD; PRT; 362 AA.
 AC P30487;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-49(B-21) B*4901 ALPHA CHAIN
 DE PRECURSOR.
 GN HLA-B OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]

| | |
|---|---|
| RESULT | 7 |
| 1B49_HUMAN | |
| ID | 1B49_HUMAN |
| AC | STANDARD; |
| PRT; | 362 AA. |
| P18464: | |
| DT | 01-NOV-1990 (Rel. 16, Created) |
| DDT | 01-NOV-1990 (Rel. 16, Last sequence update) |
| DE | 01-FEB-1996 (Rel. 33, Last annotation update) |
| HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-51(B-5) B*5101 ALPHA CHAIN PRECURSOR. | |
| GN | GLA-B OR HLAB. |
| OS | Homo sapiens (Human). |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; |
| CC | Eutheria; Primates; Catarrhini; Hominoidea; Homo. |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |

| FRAM, FRC04239, MHC-1, i. | 1 | 24 | |
|---|----------|-----------|---|
| MHC 1; Transmembrane; Glycoprotein; Signal. | 25 | 362 | |
| FT FT FT FT FT FT | 25 | 362 | |
| SIGNAL CHAIN | | | |
| FT FT | 25 | 362 | HIA CLASS I HISTOCOMPATIBILITY ANTIGEN, |
| FT FT | 25 | 114 | B-51(B-5) B*5101 ALPHA CHAIN. |
| FT FT | 115 | 206 | EXTRACELLULAR ALPHA-1. |
| FT FT | 207 | 298 | EXTRACELLULAR ALPHA-2. |
| FT FT | 299 | 308 | EXTRACELLULAR ALPHA-3. |
| FT FT | 309 | 332 | CONNECTING PEPTIDE. |
| FT FT | 309 | 332 | |
| DOMAIN | 333 | 362 | CYTOPLASMIC TAIL. |
| FT FT | 110 | 110 | BY SIMILARITY. |
| FT FT | CARBOHYD | | BY SIMILARITY. |
| FT FT | 125 | 188 | BY SIMILARITY. |
| FT FT | DISULFID | 227 | BY SIMILARITY. |
| FT FT | DISULFID | 227 | BY SIMILARITY. |
| FT FT | 362 AA; | 40566 MW; | 4D846F30 CRC32; |
| SO SEQUENCE | | | |

```

Query Match      89.8%;   Score 44;   DB 1;   Length 362;
Best Local Similarity 90.0%;   Pred. No. 0.076;
Matches 9;   Conservative 1;   Mismatches 0;   Indels 0;   Gaps 0;

QY      1 REDRLRYLY 10
         ||:|||||
Db      99 RENRLRYLY 108

```

```

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-52(B-5) B*5201 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 89080265.
RA HAYASHI H., ENNIS P.D., ARIGA H., SALTER R.D., PARHAM P., KANO K.,
RA TAKIGUCHI M.;
RT "HLA-B51 and HLA-Bw52 differ by only two amino acids which are in the
RT helical region of the alpha 1 domain.";
RL J. Immunol. 142:306-311(1989).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22799; AAA59645.1; ALT SEQ.
DR EMBL; M22793; AAA59645.1; JOINED.
DR EMBL; M22794; AAA59645.1; JOINED.
DR EMBL; M22795; AAA59645.1; JOINED.
DR EMBL; M22796; AAA59645.1; JOINED.
DR EMBL; M22797; AAA59645.1; JOINED.
DR EMBL; M22798; AAA59645.1; JOINED.
DR PIR; B30345; B30345.
DR PIR; B30348; B30348.
DR HSP; P30491; IAIM.
DR MIM; 142830; -.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; ig; 1.
DR PFAM; PF00129; MHC_I; 1.
DR MHC_I; Transmembrane; Glycoprotein; Signal.
KW SIGNAL
FT SIGNAL 1 24
FT CHAIN 25 362
FT FT
FT FT
FT DOMAIN 25 114
FT DOMAIN 115 206
FT DOMAIN 207 298
FT DOMAIN 299 308
FT TRANSMEM 309 332
FT DOMAIN 333 362
FT CARBOHYD 110 110
FT BY SIMILARITY.
FT DISULFID 125 188
FT BY SIMILARITY.
FT DISULFID 227 283
FT BY SIMILARITY.
SQ SEQUENCE 362 AA; 40521 MW; 3B436FE8 CRC32;

Query Match 89.8%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.076;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
Db 99 RENLRALRY 108
;

RESULT 10
ID 1B54_HUMAN STANDARD; PRT; 362 AA.
AC P30491;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-51(B-5) B*5104 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 92269955.
RA BELICH M.P., MADRIGAL J.A., HILDEBRAND W.H., ZEMMOUR J.,
RA WILLIAMS R.C., LUZ R., PETZL-ERLER M.L., PARHAM P.;
RT "Unusual HLA-B alleles in two tribes of Brazilian Indians.";
RL Nature 357:326-329(1992).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z15143; CAA78849.1; -.
DR HSP; P30491; IAIM.
DR MIM; 142830; -.
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; ig; 1.
DR PFAM; PF00129; MHC_I; 1.
DR MHC_I; Transmembrane; Glycoprotein; Signal.
KW SIGNAL
FT SIGNAL 1 24
FT CHAIN 25 362
FT FT
FT FT
FT DOMAIN 25 114
FT DOMAIN 115 206
FT DOMAIN 207 298
FT DOMAIN 299 308
FT TRANSMEM 309 332
FT DOMAIN 333 362
FT CARBOHYD 110 110
FT BY SIMILARITY.
FT DISULFID 125 188
FT BY SIMILARITY.
FT DISULFID 227 283
FT BY SIMILARITY.
SQ SEQUENCE 362 AA; 40560 MW; F22F08AB CRC32;

Query Match 89.8%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.076;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
Db 99 RENLRALRY 108
;

RESULT 9
ID 1B53_HUMAN STANDARD; PRT; 362 AA.
AC P30490;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)

```

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, BW-53 B*5301 ALPHA CHAIN
DE PRECURSOR.
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 91033941.
RA HAYASHI H., OORA T., NAKAYAMA S., SEKIMATA M., KANO K.,
RA TAKIGUCHI M.;
RT "Alliospecificities between HLA-B*53 and HLA-B*35 are generated by
RT substitution of the residues associated with HLA-B*4/B*6 public
RT epitopes.";
RL Immunogenetics 32:195-199(1990).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 25-302.
RX MEDLINE; 96209672.
RA SMITH K.J., REID S.W., HARLOS K., MCMICHAEL A.J., STUART D.I.,
RA BELL J.I., JONES E.Y.;
RT "Bound water structure and polymorphic amino acids act together to
RT allow the binding of different peptides to MHC class I HLA-B*53.";
RL Immunity 4:215-228(1996).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M58636; AAA36228.1; -
DR PIR; A45834; A45834.
DR PDB; 1AJM; 08-APR-98.
DR PDB; 1AJO; 08-APR-98.
DR MIM; 142830; -
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT BW-53 B*5301 ALPHA CHAIN.
FT DOMAIN 25 114
FT EXTRACELLULAR ALPHA-1.
FT DOMAIN 115 206
FT EXTRACELLULAR ALPHA-2.
FT DOMAIN 207 298
FT EXTRACELLULAR ALPHA-3.
FT DOMAIN 299 308
FT CONNECTING PEPTIDE.
FT TRANSMEM 309 332
FT CYTOPLASMIC TAIL.
FT DOMAIN 333 362
FT CARBOHYD 110 110
FT BY SIMILARITY.
FT DISULFID 125 188
FT BY SIMILARITY.
FT DISULFID 227 283
FT BY SIMILARITY.
SQ SEQUENCE 362 AA; 40495 MW; 2BDC746E CRC32;

Query Match 89.8%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.076;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
DB 99 RENLRALRY 108
|:|:|:|:|:|:|

RESULT 11
1B60_HUMAN STANDARD; PRT; 362 AA.
ID 1B60_HUMAN
AC P18465;

DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-57(B-17) B*5701 ALPHA
DE CHAIN PRECURSOR (BW57.1).
GN HLA-B OR HLAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 90207291.
RA ENNIS P.D., ZEMMOUR J., SALTER R.D., PARHAM P.;
RT "Rapid cloning of HLA-A,B CDNA by using the polymerase chain
RT reaction: frequency and nature of errors produced in amplification.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91067476.
RA ISAMAT M., GIRDLESTONE J., MILSTEIN C.;
RT "Nucleotide sequence of an HLA-B*57 gene.";
RL Nucleic Acids Res. 18:6702-6702(1990).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M32318; AAA36231.1; -
DR EMBL; X55711; CAA39244.1; -
DR PIR; S12622; S12622.
DR PIR; D35997; D35997.
DR HSP; P30491; 1AJM.
DR MIM; 142830; -
DR PROSITE; PS00290; IG_MHC; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT B-57(B-17) B*5701 ALPHA CHAIN.
FT DOMAIN 25 114
FT EXTRACELLULAR ALPHA-1.
FT DOMAIN 115 206
FT EXTRACELLULAR ALPHA-2.
FT DOMAIN 207 298
FT EXTRACELLULAR ALPHA-3.
FT DOMAIN 299 308
FT CONNECTING PEPTIDE.
FT TRANSMEM 309 332
FT CYTOPLASMIC TAIL.
FT DOMAIN 333 362
FT CARBOHYD 110 110
FT BY SIMILARITY.
FT DISULFID 125 188
FT BY SIMILARITY.
FT DISULFID 227 283
FT BY SIMILARITY.
SQ SEQUENCE 362 AA; 40224 MW; D91DF8DD CRC32;

Query Match 89.8%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.076;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
DB 99 RENLRALRY 108
|:|:|:|:|:|:|

RESULT 12
1B61_HUMAN STANDARD; PRT; 362 AA.
ID 1B61_HUMAN
AC P30497;

DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B-57(B-17) B*5702 ALPHA CHAIN
 DE PRECURSOR (B*57.2).
 GN HLA-B OR HLAB.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 93056508.
 RA MADRIGAL J.A., ZELITCH M.P., HILDBRAND W.H., BENJAMIN R.J.,
 RA LITTLE A.-M., BEMOUR J., ENNIS P.D., WARD F.E., PETZL-ERLER M.L.,
 RA MARTELL R.W., DU TOIT E.D., PARHAM P.;
 RT "Distinctive HLA-A-B antigens of black populations formed by
 RT interallelic conversion.";
 RL J. Immunol. 149:3411-3415(1992).
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X61707; CAA43876.1; -.
 DR PIR; S16774; S16774.
 DR HSP; P30491; IALM.
 DR MIN; 142830; -.
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; Ig; 1.
 DR PFAM; PF00129; MHC_I; 1.
 DR MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT DOMAIN 25 114 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
 FT DOMAIN 115 206 B-57(B-17) B*5702 ALPHA CHAIN.
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-1.
 FT DOMAIN 299 308 EXTRACELLULAR ALPHA-2.
 FT TRANSMEM 309 332 EXTRACELLULAR ALPHA-3.
 FT DOMAIN 333 362 CONNECTING PEPTIDE.
 FT CARBOHYD 110 110 CYTOPLASMIC TAIL.
 FT DISULFID 125 188 BY SIMILARITY.
 FT DISULFID 227 283 BY SIMILARITY.
 SQ SEQUENCE 362 AA; 40342 MW; 628C2156 CRC32;
 Query Match 89.8%; Score 44; DB 1; Length 362;
 Best Local Similarity 90.0%; Pred. No. 0.076;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 REDLRALRY 10
 Db 99 RENLRALRY 108
 RESULT 13
 ID 1B62_HUMAN STANDARD; PRT; 362 AA.
 AC P10319;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B*58(B-17) B*5801 ALPHA
 DE CHAIN PRECURSOR.
 GN HLA-B OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 93056508.
 RA MADRIGAL J.A., ZELITCH M.P., HILDBRAND W.H., BENJAMIN R.J.,
 RA LITTLE A.-M., BEMOUR J., ENNIS P.D., WARD F.E., PETZL-ERLER M.L.,
 RA MARTELL R.W., DU TOIT E.D., PARHAM P.;
 RT "Distinctive HLA-A-B antigens of black populations formed by
 RT interallelic conversion.";
 RL J. Immunol. 149:3411-3415(1992).
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X61707; CAA43876.1; -.
 DR PIR; S16774; S16774.
 DR HSP; P30491; IALM.
 DR MIN; 142830; -.
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; Ig; 1.
 DR PFAM; PF00129; MHC_I; 1.
 DR MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT DOMAIN 25 114 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
 FT DOMAIN 115 206 B-57(B-17) B*5702 ALPHA CHAIN.
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-1.
 FT DOMAIN 299 308 EXTRACELLULAR ALPHA-2.
 FT TRANSMEM 309 332 EXTRACELLULAR ALPHA-3.
 FT DOMAIN 333 362 CONNECTING PEPTIDE.
 FT CARBOHYD 110 110 CYTOPLASMIC TAIL.
 FT DISULFID 125 188 BY SIMILARITY.
 FT DISULFID 227 283 BY SIMILARITY.
 SQ SEQUENCE 362 AA; 40342 MW; 628C2156 CRC32;
 Query Match 89.8%; Score 44; DB 1; Length 362;
 Best Local Similarity 90.0%; Pred. No. 0.076;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 REDLRALRY 10
 Db 99 RENLRALRY 108
 RESULT 13
 ID 1B62_HUMAN STANDARD; PRT; 362 AA.
 AC P10319;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, B*58(B-17) B*5801 ALPHA
 DE CHAIN PRECURSOR.
 GN HLA-B OR HLAB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86008247.
 RA WAYS J.P., COPPIN H.L., PARHAM P.;
 RT "The complete primary structure of HLA-B*58.";
 RL J. Biol. Chem. 260:11924-11933(1985).
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M11799; AAA59628.1; -.
 DR EMBL; AB008102; BAA22916.1; -.
 DR PIR; A23895; HLHUB8.
 DR HSP; P30491; IALM.
 DR MIN; 142830; -.
 DR PROSITE; PS00290; IG_MHC; 1.
 DR PFAM; PF00047; Ig; 1.
 DR PFAM; PF00129; MHC_I; 1.
 DR MHC I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 362
 FT DOMAIN 25 114 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
 FT DOMAIN 115 206 B*5801 ALPHA CHAIN.
 FT DOMAIN 207 298 EXTRACELLULAR ALPHA-1.
 FT DOMAIN 299 308 EXTRACELLULAR ALPHA-2.
 FT TRANSMEM 309 332 EXTRACELLULAR ALPHA-3.
 FT DOMAIN 333 362 CONNECTING PEPTIDE.
 FT CARBOHYD 110 110 CYTOPLASMIC TAIL.
 FT DISULFID 125 188 BY SIMILARITY.
 FT DISULFID 227 283 BY SIMILARITY.
 SQ SEQUENCE 362 AA; 40337 MW; 3E5E7534 CRC32;
 Query Match 89.8%; Score 44; DB 1; Length 362;
 Best Local Similarity 90.0%; Pred. No. 0.076;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 REDLRALRY 10
 Db 99 RENLRALRY 108
 RESULT 14
 ID 1B62_HUMAN STANDARD; PRT; 362 AA.
 AC P01893;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-DEC-1999 (Rel. 39, Last annotation update)
 DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, ALPHA CHAIN H PRECURSOR
 DE (HLA-AR) (HLA-12.4).
 GN HLA-H OR HLAH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

RN  SEQUENCE FROM N.A.
RX  MEDLINE: 82151002.
RT  MALISSEN M., MALISSEN B., JORDAN B.R.;
RA  "Exon/intron organization and complete nucleotide sequence of an HLA
RL  gene.";
RT  Proc. Natl. Acad. Sci. U.S.A. 79:893-897(1982).
CC  -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC  THE IMMUNE SYSTEM. COULD BE THE PRODUCT OF A PSEUDOGENE.
CC  -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC  MICROGLOBULIN).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: J00191; AAA36218.1; ALT_INIT.
DR  PIR: A02189; HLH12.
DR  HSP: P03989; LHS4.
DR  MIM: 142800; -.
DR  PROSITE: PS00290; IG_MHC; 1.
DR  PFAM: PF00047; ig; 1.
DR  PFAM: PF00129; MHC_I; 1.
DR  MHC I; Transmembrane; Glycoprotein; Signal.
KW  SIGNAL
FT  CHAIN 1 24
FT  CHAIN 25 362
FT  HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT  ALPHA CHAIN H.
FT  DOMAIN 25 114
FT  DOMAIN 115 206
FT  DOMAIN 207 298
FT  DOMAIN 299 308
FT  DOMAIN 309 332
FT  TRANSMEM 333 362
FT  CARBOHYD 110 110
FT  DISULFID 227 283
FT  BY SIMILARITY.
SQ  SEQUENCE 362 AA; 40850 MW; 5E610F63 CRC32;

Query Match 89.8%; Score 44; DB 1; Length 362;
Best Local Similarity 90.0%; Pred. No. 0.076;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
DB 99 RENDLRALRY 108
||:|||||

RESULT 15
ID 1A23_HUMAN STANDARD; PRT: 365 AA.
AC P30447;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-23(A-9) ALPHA CHAIN
DE PRECURSOR.
DE
GN HLA-A OR HLAA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A. (A*2301).
RX MEDLINE: 92104637.
RA LITTLE A.-M., MADRIGAL J.A., PARHAM P.;
RT "Molecular definition of an elusive third HLA-A9 molecule: HLA-A9.3.";
RL Immunogenetics 35:41-45(1992).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-

```

```

CC  MICROGLOBULIN).
CC  -!- POLYMORPHISM: THE ONLY ALLELE OF A-23 KNOWN IS A*2301 WHICH IS
CC  SHOWN HERE.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: M64742; AAA03662.1; -.
DR  HSP: P01892; LAQD.
DR  MIM: 142800; -.
DR  PROSITE: PS00290; IG_MHC; 1.
DR  PFAM: PF00047; ig; 1.
DR  PFAM: PF00129; MHC_I; 1.
DR  MHC I; Transmembrane; Glycoprotein; Signal.
KW  SIGNAL
FT  CHAIN 1 24
FT  CHAIN 25 365
FT  HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT  A-23(A-9) ALPHA CHAIN.
FT  DOMAIN 25 114
FT  DOMAIN 115 206
FT  DOMAIN 207 298
FT  DOMAIN 299 308
FT  DOMAIN 309 332
FT  TRANSMEM 333 365
FT  CARBOHYD 110 110
FT  DISULFID 125 188
FT  BY SIMILARITY.
SQ  SEQUENCE 365 AA; 40732 MW; B1C21094 CRC32;

Query Match 89.8%; Score 44; DB 1; Length 365;
Best Local Similarity 90.0%; Pred. No. 0.076;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
DB 99 RENDLRALRY 108
||:|||||

Search completed: February 8, 2000, 00:59:52
Job time: 3781 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2000, 13:17:39 ; Search time 209.03 Seconds
(without alignments)
3.317 Million cell updates/sec

Title: US-08-653-294-12
Perfect score: 49
Sequence: 1 REDLRALRY 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SPTREMBL12:.*
1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mhc:.*
8: sp_organelle:.*
9: sp_phage:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 44 | 89.8 | 89 | 7 | O19569 homo sapien |
| 2 | 44 | 89.8 | 90 | 7 | O46697 gorilla gor |
| 3 | 44 | 89.8 | 133 | 7 | O19189 homo sapien |
| 4 | 44 | 89.8 | 138 | 7 | O78209 homo sapien |
| 5 | 44 | 89.8 | 172 | 7 | O19770 homo sapien |
| 6 | 44 | 89.8 | 172 | 7 | O19774 homo sapien |
| 7 | 44 | 89.8 | 172 | 7 | O19775 homo sapien |
| 8 | 44 | 89.8 | 172 | 7 | O19780 homo sapien |
| 9 | 44 | 89.8 | 172 | 7 | O95364 homo sapien |
| 10 | 44 | 89.8 | 172 | 7 | O19771 homo sapien |
| 11 | 44 | 89.8 | 172 | 7 | O19772 homo sapien |
| 12 | 44 | 89.8 | 172 | 7 | O19773 homo sapien |
| 13 | 44 | 89.8 | 175 | 7 | O29694 homo sapien |
| 14 | 44 | 89.8 | 180 | 7 | O19607 homo sapien |
| 15 | 44 | 89.8 | 180 | 7 | O19608 homo sapien |
| 16 | 44 | 89.8 | 180 | 7 | O19609 homo sapien |
| 17 | 44 | 89.8 | 180 | 7 | O19610 homo sapien |
| 18 | 44 | 89.8 | 180 | 7 | O19611 homo sapien |
| 19 | 44 | 89.8 | 180 | 7 | O19612 homo sapien |
| 20 | 44 | 89.8 | 180 | 7 | O19613 homo sapien |

| | | | | | | |
|----|----|------|-----|---|--------|--------------------|
| 21 | 44 | 89.8 | 181 | 7 | O46703 | 046703 homo sapien |
| 22 | 44 | 89.8 | 181 | 7 | O62917 | O62917 homo sapien |
| 23 | 44 | 89.8 | 181 | 7 | O62892 | O62892 homo sapien |
| 24 | 44 | 89.8 | 181 | 7 | O62899 | O62899 homo sapien |
| 25 | 44 | 89.8 | 181 | 7 | O62920 | O62920 homo sapien |
| 26 | 44 | 89.8 | 181 | 7 | O62922 | O62922 homo sapien |
| 27 | 44 | 89.8 | 181 | 7 | O62923 | O62923 homo sapien |
| 28 | 44 | 89.8 | 181 | 7 | O19623 | O19623 homo sapien |
| 29 | 44 | 89.8 | 181 | 7 | O19747 | O19747 homo sapien |
| 30 | 44 | 89.8 | 181 | 7 | O29667 | O29667 homo sapien |
| 31 | 44 | 89.8 | 181 | 7 | O30198 | O30198 homo sapien |
| 32 | 44 | 89.8 | 181 | 7 | O29708 | O29708 homo sapien |
| 33 | 44 | 89.8 | 181 | 7 | O19631 | O19631 homo sapien |
| 34 | 44 | 89.8 | 181 | 7 | O19769 | O19769 homo sapien |
| 35 | 44 | 89.8 | 181 | 7 | O29724 | O29724 homo sapien |
| 36 | 44 | 89.8 | 181 | 7 | O29910 | O29910 homo sapien |
| 37 | 44 | 89.8 | 181 | 7 | P79559 | P79559 homo sapien |
| 38 | 44 | 89.8 | 181 | 7 | O29679 | O29679 homo sapien |
| 39 | 44 | 89.8 | 181 | 7 | O19521 | O19521 homo sapien |
| 40 | 44 | 89.8 | 181 | 7 | O19597 | O19597 homo sapien |
| 41 | 44 | 89.8 | 181 | 7 | O29909 | O29909 homo sapien |
| 42 | 44 | 89.8 | 181 | 7 | O29701 | O29701 homo sapien |
| 43 | 44 | 89.8 | 181 | 7 | O29841 | O29841 homo sapien |
| 44 | 44 | 89.8 | 181 | 7 | O19354 | O19354 gorilla gor |
| 45 | 44 | 89.8 | 181 | 7 | O29765 | O29765 homo sapien |

ALIGNMENTS

RESULT 1
O19569 PRELIMINARY; PRT; 89 AA.
AC O19569;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE MHC CLASS I ANTIGEN (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CAO K., BURDET L., ZHANG G., FERNANDEZ-VINA M.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF017320; AAB70286.2; -
KW MHC.
FT NON_TER 1 1
FT NON_TER 89 89
SQ SEQUENCE 89 AA; 10606 MW; 99D11089 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 89;
Best Local Similarity 90.0%; Pred. No. 0.14;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
Db 74 REDLRALRY 83

RESULT 2
O46697 PRELIMINARY; PRT; 90 AA.
AC O46697;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE MHC CLASS I ANTIGEN HLA-H ORTHOLOG (FRAGMENT).
GN HLA-H.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SHAMBA;
 RA GRIMSLEY C., MATHER K.A., OBER C.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF022172; AAC99794.1; -
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1
 FT NON_TER 90
 FT NON_TER 90
 SQ SEQUENCE 90 AA; 10689 MW; 5E5F2495 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 90;
 Best Local Similarity 90.0%; Pred. No. 0.15;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDLRALRY 10
 ||:|||||
 Db 75 RENLRALRY 84

RESULT 3

OL19189 PRELIMINARY; PRT; 133 AA.
 AC OL19189;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE MHC CLASS I HISTOCOMPATIBILITY ANTIGEN-B (HLA-B-27KSH) (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LEUKOCYTE;
 RA PETERSDORF E.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U18659; AAB60357.1; -
 DR MIM; 142830; -
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC I.
 FT NON_TER 1
 FT NON_TER 133
 FT NON_TER 133
 SQ SEQUENCE 133 AA; 15491 MW; 3A3BC802 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 133;
 Best Local Similarity 90.0%; Pred. No. 0.22;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDLRALRY 10
 ||:|||||
 Db 27 RENLRALRY 36

RESULT 4

OL78209 PRELIMINARY; PRT; 138 AA.
 AC OL78209;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
 DE HUMAN LEUKOCYTE ANTIGEN PRECURSOR (FRAGMENT).
 GN HLA-A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 98007772.

RA LAFORET M., FROELICH N., PARISSIAS A., BAUSINGER H., PFEIFFER B.,
 RA TONGIO M.M.;
 RT "An intronic mutation responsible for a low level of expression of an
 RT HLA-A*24 allele."
 RL Tissue Antigens 50:340-346(1997).
 DR EMBL; Z7423; CAA96533.1; -
 DR PFAM; PF00129; MHC_I; 1.
 KW Signal; MHC.
 FT SIGNAL 1
 FT NON_TER 138
 FT NON_TER 138
 SQ SEQUENCE 138 AA; 15610 MW; B8417FA0 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 138;
 Best Local Similarity 90.0%; Pred. No. 0.23;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDLRALRY 10
 ||:|||||
 Db 99 RENLRALRY 108

RESULT 5

OL19770 PRELIMINARY; PRT; 172 AA.
 AC OL19770;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE MHC CLASS I HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGRONG E.,
 RA BEJCHANDRA S., JUJI T., TOKUNAGA K.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U90421; AAB50144.1; -
 DR EMBL; U90420; AAB50144.1; JOINED.
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1
 FT NON_TER 172
 FT NON_TER 172
 SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 172;
 Best Local Similarity 90.0%; Pred. No. 0.3;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REDLRALRY 10
 ||:|||||
 Db 68 RENLRALRY 77

RESULT 6

OL19774 PRELIMINARY; PRT; 172 AA.
 AC OL19774;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE MHC CLASS I HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGRONG E.,
 RA BEJCHANDRA S., BLASZYK R., GROSSE-WILDE H.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90423; AAB50145.1; -.
DR EMBL; U90422; AAB50145.1; JOINED.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 172;
Best Local Similarity 90.0%; Pred. No. 0.3;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
DB 68 RENLRALRY 77

RESULT 7
O19775 PRELIMINARY; PRT; 172 AA.

ID O19775
AC O19775;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)

DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]

RP SEQUENCE FROM N.A.
RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,
RA BEJCHANDRA S.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90425; AAB50146.1; -.
DR EMBL; U90424; AAB50146.1; JOINED.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.

FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 172;
Best Local Similarity 90.0%; Pred. No. 0.3;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
DB 68 RENLRALRY 77

RESULT 8
O19780 PRELIMINARY; PRT; 172 AA.

ID O19780
AC O19780;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)

DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]

RP SEQUENCE FROM N.A.
RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,
RA BEJCHANDRA S., JUJI T., TOKUNAGA K.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90419; AAB50143.1; -.
DR EMBL; U90418; AAB50143.1; JOINED.

QY 1 REDLRALRY 10
DB 68 RENLRALRY 77

RESULT 10
O19771 PRELIMINARY; PRT; 172 AA.

ID O19771
AC O19771;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)

DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]

RP SEQUENCE FROM N.A.
RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D., RUNGROUNG E.,
RA SIRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90612; AAB50151.1; -.
DR EMBL; U90611; AAB50151.1; JOINED.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.

FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 20026 MW; 4D9A1043 CRC32;

DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 19909 MW; CAAE5641 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 172;
Best Local Similarity 90.0%; Pred. No. 0.3;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
DB 68 RENLRALRY 77

RESULT 9
O95364 PRELIMINARY; PRT; 172 AA.

ID O95364
AC O95364;
DT 01-FEB-1997 (TEMBLrel. 02, Created)
DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)

DE MHC HLA-B*51 PROTEIN (FRAGMENT).
GN HLA-B*51pA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]

RP SEQUENCE FROM N.A.
RA BUASCIK R.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X96473; CAA65327.1; -.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.

FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 19942 MW; 1A73E47D CRC32;

Query Match 89.8%; Score 44; DB 7; Length 172;
Best Local Similarity 90.0%; Pred. No. 0.3;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
DB 65 RENLRALRY 74

RESULT 10
O19771 PRELIMINARY; PRT; 172 AA.

ID O19771
AC O19771;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)

DE MHC CLASS I HLA-B (FRAGMENT).
GN HLA-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]

RP SEQUENCE FROM N.A.
RA CHANDANAYINGYONG D., SIRIKONG M., LONGTA K., SRINAK D.,
RA SIRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90612; AAB50151.1; -.
DR EMBL; U90611; AAB50151.1; JOINED.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.

FT NON_TER 1 1
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 20026 MW; 4D9A1043 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 172;
 Best Local Similarity 90.0%; Pred. No. 0.3;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
 ||:|||||||
 Db 68 RENLRALRY 77

RESULT 11
 Q19772 PRELIMINARY; PRT: 172 AA.
 AC Q19772;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE MHC CLASS I HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D.,
 RA SIRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U90614; AAB50244.1; -
 DR EMBL; U90613; AAB50244.1; JOINED.
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 172 172
 SQ SEQUENCE 172 AA; 20026 MW; 4D9A1043 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 172;
 Best Local Similarity 90.0%; Pred. No. 0.3;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
 ||:|||||||
 Db 68 RENLRALRY 77

RESULT 12
 Q19773 PRELIMINARY; PRT: 172 AA.
 AC Q19773;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE MHC CLASS I HLA-B (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYONG D., SIRIKONG M., LONGTA K., SRINAK D.,
 RA SIRIBOONRIT U., RUNGROUNG E., BEJCHANDRA S.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U90616; AAB50245.1; -
 DR EMBL; U90615; AAB50245.1; JOINED.
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 172 172
 SQ SEQUENCE 172 AA; 20052 MW; F6214671 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 172;

Best Local Similarity 90.0%; Pred. No. 0.3;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
 ||:|||||||
 Db 68 RENLRALRY 77

RESULT 13
 Q29694 PRELIMINARY; PRT: 175 AA.
 ID Q29694;
 AC Q29694;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE MHC CLASS I HLA-B ANTIGEN (FRAGMENT).
 GN HLA-B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD;
 RA PETERSDORF E.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U28759; AAB60367.1; -
 DR HSSP; P10318; 1ROG.
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT VARIANT 67 67 Y -> D.
 FT VARIANT 73 73 I -> T.
 FT NON_TER 175 175
 SQ SEQUENCE 175 AA; 20332 MW; 83A0C5C3 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 175;
 Best Local Similarity 90.0%; Pred. No. 0.3;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 REDLRALRY 10
 ||:|||||||
 Db 68 RENLRALRY 77

RESULT 14
 Q19607 PRELIMINARY; PRT: 180 AA.
 ID Q19607;
 AC Q19607;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE MHC CLASS I HLA-A (FRAGMENT).
 GN HLA-A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CHANDANAYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
 RA RUNGROUNG E., BEJCHANDRA S.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF030920; AAB87056.1; -
 DR EMBL; AF030919; AAB87056.1; JOINED.
 DR HSSP; P01891; 1TMC.
 DR PFAM; PF00129; MHC_I; 1.
 KW MHC.
 FT NON_TER 1 1
 FT NON_TER 180 180
 SQ SEQUENCE 180 AA; 20811 MW; CECC3537 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 180;

Best Local Similarity 90.0%; Pred. No. 0.31; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 1;

QY 1 REDLRALRY 10
II:IIIIII
Db 74 RNLRLALRY 83

RESULT 15

O19608 PRELIMINARY; PRT; 180 AA.
AC O19608:
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE MHC CLASS I HLA-A (FRAGMENT).
GN HLA-A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHANDANAYONG D., SIRIKONG M., LUANGTRAKOOL K., SRINAK D.,
RA RUNGROUNG E., BEJCHANDRA S.:
RL Submitted (Oct-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF030922; AAB87057.1; -.
DR EMBL; AF030921; AAB87057.1; JOINED.
DR HSSP; P01891; 1TMC.
DR PFAM; PF00129; MHC_I; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 180 180
SQ SEQUENCE 180 AA; 20811 MW; CECC3537 CRC32;

Query Match 89.8%; Score 44; DB 7; Length 180;

Best Local Similarity 90.0%; Pred. No. 0.31; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 1;

QY 1 REDLRALRY 10
II:IIIIII
Db 74 RNLRLALRY 83

Search completed: February 8, 2000, 13:17:39
Job time: 32488 sec

THIS PAGE BLANK (USPTO)

| Strd | Orig | ZScore | EScore | Len | Documentation | ... |
|-------------------|------|--------|--------|-----|---------------|-----------------------------|
| 9b_pr1:HSULABHBA | + | 44.00 | 0.1647 | 250 | I Y08692 | H.sapiens HLA-B gene, ex |
| 9b_pr1:HSULABHBB | + | 44.00 | 0.1647 | 250 | I Y08693 | H.sapiens HLA-B gene, ex |
| 9b_pr1:HSULABHBD | + | 44.00 | 0.1647 | 250 | I Y08694 | H.sapiens HLA-B gene, ex |
| 9b_pr1:HSULABAB1 | + | 44.00 | 0.1647 | 250 | I Y09420 | Human cell line THAI DCH |
| 9b_pr2:HSULABD1 | + | 44.00 | 0.1647 | 250 | I Y09611 | Human cell line THAI DCH |
| 9b_pr2:HSULABD1 | + | 44.00 | 0.1647 | 250 | I Y09611 | Human cell line THAI DCH |
| 9b_pr2:HSULABG1 | + | 44.00 | 0.1647 | 250 | I Y09615 | Human cell line THAI DCH |
| 9b_pr2:HSULABJ1 | + | 44.00 | 0.1647 | 250 | I Y09422 | Human cell line THAI DCH |
| 9b_pr2:HSULABJ1 | + | 44.00 | 0.1647 | 250 | I Y09424 | Human cell line THAI DCH |
| 9b_pr2:HSULABT1 | + | 44.00 | 0.1647 | 250 | I Y09418 | Human cell line THAI DCH |
| 9b_pr4:AF022159 | + | 44.00 | 0.1793 | 255 | I AF022159 | Homo sapiens isolate 0 |
| 9b_pr4:AF022160 | + | 44.00 | 0.1713 | 259 | I AF022160 | Homo sapiens isolate 0 |
| 9b_pat:114590 | + | 44.00 | 0.1708 | 270 | I I4590 | Sequence 67 from patent |
| 9b_pat:114591 | + | 44.00 | 0.1708 | 270 | I I4591 | Sequence 68 from patent |
| 9b_pat:114592 | + | 44.00 | 0.1793 | 270 | I I4592 | Sequence 69 from patent |
| 9b_pat:114593 | + | 44.00 | 0.1708 | 270 | I I4593 | Sequence 70 from patent |
| 9b_pr1:HSB51EX2 | + | 44.00 | 0.1708 | 270 | I J11228 | H.sapiens HLA-B*51 gene, al |
| 9b_pr1:HSHLAAL1 | + | 44.00 | 0.1793 | 270 | I X81161 | H.sapiens HLA-A*1 gene, al |
| 9b_pr1:NUMHB512 | + | 44.00 | 0.1708 | 270 | I M22787 | Human MHC class I HLA-B*5 |
| 9b_pr1:NUMHBW2 | + | 44.00 | 0.1708 | 270 | I M22794 | Human MHC class I HLA-B*W |
| 9b_pr2:1HMCCHU1A | + | 44.00 | 0.1708 | 270 | I U63595 | Human MHC class I antigen |
| 9b_pr2:HSAA011699 | + | 44.00 | 0.1793 | 270 | I AJ011699 | Homo sapiens HLA-A*1 gene |
| 9b_pr2:HSAA133780 | + | 44.00 | 0.1708 | 270 | I AJ133780 | Homo sapiens HLA-B*1 gene |
| 9b_pr2:HSAA238971 | + | 44.00 | 0.1793 | 270 | I AJ238971 | Homo sapiens HLA-B*1 gene |
| 9b_pr2:HSAA239035 | + | 44.00 | 0.1708 | 270 | I AJ239035 | Homo sapiens HLA-A*1 gene |
| 9b_pr2:HSHL24JV01 | + | 44.00 | 0.1708 | 270 | I U37110 | Human HLA-A*24 gene, alle |
| 9b_pr2:HSHL24SA01 | + | 44.00 | 0.1793 | 270 | I U37114 | Human HLA-A*24 gene, alle |
| 9b_pr2:HSHL24YM01 | + | 44.00 | 0.1708 | 270 | I U37112 | Human HLA-A*24 gene, alle |
| 9b_pr2:HSHLA24S1 | + | 44.00 | 0.1793 | 270 | I U19887 | Human MHC class I antigen |
| 9b_pr2:HSHLABEXN1 | + | 44.00 | 0.1708 | 270 | I U76400 | Human HLA-B*1 gene, alle |
| 9b_pr2:1UMMHHLAAF | + | 44.00 | 0.1793 | 270 | I L43530 | Homo sapiens (clone K920) |
| 9b_pr2:1A3528 | + | 44.00 | 0.1708 | 270 | I L43528 | Homo sapiens (clone K620) |
| 9b_pr2:X83402 | + | 44.00 | 0.1793 | 270 | I X83402 | H.sapiens HLA-B*1 gene, ex |
| 9b_pr3:HA1109HLA1 | + | 44.00 | 0.1793 | 270 | I AF030923 | Homo sapiens MHC class I |
| 9b_pr3:HS1109HLA1 | + | 44.00 | 0.1708 | 270 | I AF030931 | Homo sapiens MHC class I |
| 9b_pr3:HS2402HLA1 | + | 44.00 | 0.1793 | 270 | I AF067436 | Homo sapiens MHC class I |
| 9b_pr3:HS507HLA1 | + | 44.00 | 0.1708 | 270 | I AF030921 | Homo sapiens MHC class I |
| 9b_pr3:HS1109HLA1 | + | 44.00 | 0.1793 | 270 | I U52815 | Homo sapiens class I anti |
| 9b_pr3:HS222HLA1 | + | 44.00 | 0.1793 | 270 | I AF030919 | Homo sapiens MHC class I |
| 9b_pr3:HS5342HLA1 | + | 44.00 | 0.1708 | 270 | I AF030925 | Homo sapiens MHC class I |
| 9b_pr3:HS538HLA1 | + | 44.00 | 0.1793 | 270 | I AF030911 | Homo sapiens MHC class I |
| 9b_pr3:HS5802HLA1 | + | 44.00 | 0.1708 | 270 | I U52813 | Human MHC class I antigen |

```

SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 250)
AUTHORS    Rojas-Munoz,A., Mendez,I. and Yunis,I.
TITLE      Molecular evolution of HLA-B locus in a small population amerindian
            community :The Nukak-Maku
JOURNAL    Unpublished
REFERENCE   2 (bases 1 to 250)
AUTHORS    Rojas-Munoz,A.
TITLE      Direct Submission
JOURNAL    Submitted (07-OCT-1996) A. Rojas-Munoz, National Institute Of
            Health, Immunogenetics, Avenida El Dorado, Carrera 50, Santafe De
            Bogota / Zona 6, COLOMBIA
FEATURES    Location/Qualifiers
            1..250
            /organism="Homo sapiens"
            /isolate="Isabel-26"
            /isolate="from amerindian community Nukak-Maku"
            /db_xref="taxon:9606"
            /chromosome="6"
            /dev_stage="adult"
            /tissue_type="blood"
            /cell_type="white"
            /lab_host="E.coli TGI"
            /clone="CHBC2(b)"
            /clone="CHBC3(b)"
            /clone="CHBC4(b)"
            14..250
            /gene="HLA-B"
            <14..>250
            /gene="HLA-B"
            /note="allele HB(b)"
            /number=2
BASE COUNT 51 a      78 c      87 g      34 t
ORIGIN
alignment_scores:
    Quality: 44.00      Length: 10
    Ratio: 4.400      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 90.000
alignment_block:
    US-08-653-294-12 x HSHLABHB ..
    Align seg 1/1 to: HSHLABHB from: 1 to: 250
    1 ArgGluAspLeuArgIleAlaLeuArgTyr 10
      |||||:::|||||
    209 CGAGAGAACTGGGATCGCTCCGCTAC 238
seq_name: gb_pr1:HSHLABHB
seq_documentation_block:
    LOCUS      HSHLABHB      250 bp      DNA      PRI      10-OCT-1996
    DEFINITION Human cell line THAI DCH010 MHC class I HLA-B gene (allele.
    ACCESSION  Y08694
    VERSION    Y08694.1 GI:1619289
    KEYWORDS   HLA-B gene; human leukocyte antigen.
    SOURCE     human.
    ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 250)
AUTHORS    Rojas-Munoz,A., Mendez,I. and Yunis,I.
TITLE      Molecular evolution of HLA-B locus in a small population amerindian
            community :The Nukak-Maku
JOURNAL    Unpublished
REFERENCE   2 (bases 1 to 250)
AUTHORS    Rojas-Munoz,A.
TITLE      Direct Submission

```

```

JOURNAL    Submitted (07-OCT-1996) A. Rojas-Munoz, National Institute Of
            Health, Immunogenetics, Avenida El Dorado, Carrera 50, Santafe De
            Bogota / Zona 6, COLOMBIA
FEATURES    Location/Qualifiers
            1..250
            /organism="Homo sapiens"
            /isolate="Norman-51"
            /isolate="from amerindian community Nukak-Maku"
            /db_xref="taxon:9606"
            /chromosome="6"
            /dev_stage="adult"
            /tissue_type="blood"
            /cell_type="white"
            /lab_host="E.coli TGI"
            /clone="CHBC1(d)"
            14..250
            /gene="HLA-B"
            <14..>250
            /gene="HLA-B"
            /note="allele HB(d)"
            /number=2
BASE COUNT 58 a      78 c      79 g      35 t
ORIGIN
alignment_scores:
    Quality: 44.00      Length: 10
    Ratio: 4.400      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 90.000
alignment_block:
    US-08-653-294-12 x HSHLABHD ..
    Align seg 1/1 to: HSHLABHD from: 1 to: 250
    1 ArgGluAspLeuArgIleAlaLeuArgTyr 10
      |||||:::|||||
    209 CGAGAGAACTGGGATCGCTCCGCTAC 238
seq_name: gb_pr2:HSHLABB1
seq_documentation_block:
    LOCUS      HSHLABB1      250 bp      DNA      PRI      22-MAR-1997
    DEFINITION Human cell line THAI DCH010 MHC class I HLA-B gene (allele
    ACCESSION  U90420
    VERSION    U90420.1 GI:1905830
    KEYWORDS   HLA-B*1513), exon 2.
    SEGMENT    1 of 2
    SOURCE     human.
    ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 250)
AUTHORS    Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
            Rungroung,E., Bejchandra,S., Juji,T. and Tokunaga,K.
            B15 alleles (B*1513)
            Unpublished
            2 (bases 1 to 250)
            Chandanayingyong,D., Sirikong,M., Longta,K., Srinak,D.,
            Rungroung,E., Bejchandra,S., Juji,T. and Tokunaga,K.
            Direct Submission
            Submitted (23-FEB-1997) Transfusion Medicine, Faculty of Medicine,
            Siriraj Hospital, Mahidol University, Prannok Road, Bangkok 10700,
            Thailand
            Location/Qualifiers
            1..250
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="VI"
            /map="6p21"
            /cell_type="lymphoblastoid"
            /cell_line="THAI DCH010"

```